



Durham E-Theses

The metaphysics of mental causation

Gibb, Sophie Catherine

How to cite:

Gibb, Sophie Catherine (2002) *The metaphysics of mental causation*, Durham theses, Durham University.
Available at Durham E-Theses Online: <http://etheses.dur.ac.uk/3863/>

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

The Metaphysics of Mental Causation

Sophie Catherine Gibb

The copyright of this thesis rests with the author.
No quotation from it should be published without
his prior written consent and information derived
from it should be acknowledged.



14 APR 2003

University of Durham

Department of Philosophy

2002

Submitted for the degree of Ph.D.

The Metaphysics of Mental Causation

Sophie Catherine Gibb

Abstract

This thesis argues that the fundamental issues within the mental causation debate are metaphysical ones. Consequently, it is only with metaphysical clarity, that any clarity can be gained in the mental causation debate. In order to provide a successful theory of mental causation one cannot divorce oneself from metaphysics. Neither can one hope to provide a theory of mental causation that is somehow neutral between the various metaphysical systems. Rather, to be plausible, a theory of mental causation must be based within an independently plausible metaphysical framework.

I divide the metaphysical issues that are of importance to the mental causation debate into three broad groups. Firstly, what causation is a relation between. Secondly, what the existence and identity conditions for properties are. Thirdly, what the causal relation is.

Part One of this thesis is concerned with the first of these issues. The interpretation of the argument from causal overdetermination, and the possible responses to it, depend upon what causation is a relation between. A belief to the contrary, has led to implausible theories of mental causation and the misrepresentation of those positions within the mental causation debate that are ontologically serious.

Part Two is concerned with property analysis. It is suggested that a plausible analysis of properties reveals that the true contenders within the mental causation debate are psychophysical reductionism on the one hand, and interactive mentalism on the other.

Part Three is concerned with the causal relation. It is argued that the mental causation debate is affected by what one understands causation to be. In particular, whether a causal closure principle that is strong enough to allow one to advance physicalism can plausibly be advanced, depends upon the theory of causation in which one is embedding psychophysical causation.

University of Durham
Department of Philosophy

2002

Declaration

I confirm that no part of the material has been submitted by me for a degree in this or any other university. No material has been generated by joint work. In all cases material from the work of others has been acknowledged and quotations and paraphrases suitably indicated.

The copyright of this thesis rests with the author. No quotation from it should be published without their prior written consent and information derived from it should be acknowledged.

Sophie Catherine Gibb

Acknowledgements

In writing this thesis I am greatly indebted to Jonathan Lowe, who has been a constant source of encouragement and inspiration. I would also like to thank my parents, Kate and Michael Gibb, for their loving care and constant support. I am also indebted to Denis Copcutt for his help and endless patience. Finally, I would like to thank those members of the Durham Philosophy Department whom, through their care and interest, have made the writing of this thesis enjoyable.

Table of Contents

Notation

vii

Introduction

1 The Problem of Mental Causation	1
2 Metaphysics and the Problem of Mental Causation	8

Part One

The Causal Relata

1 Mental Causation and the Causal Relata	13
1.1 Preliminaries	15
1.2 The Particularity of the Causal Relata	16
1.3 Causation as a Relation between Events	17
2 Davidson's Anomalous Monism and Property Epiphenomenalism	20
2.1 Anomalous Monism	20
2.2 Property Epiphenomenalism	22
2.3 Why Davidson is not a Type Epiphenomenalist	25
2.4 Predicate Epiphenomenalism	29
2.5 A Rejection of Davidson's Theory of the Causal Relata	31
3 Coarse-grained Theories of the Causal Relata	36
3.1 Coarse-grained Theories and the Qualitative Specificity of Causation	36
3.2 The Transitivity of Causation and the Rejection of Coarse-grained Theories	38
3.3 A Re-articulation of the Argument from Causal Overdetermination	41
4 The Properties of Causation and the Mental Causation Debate	43
4.1 Property-instantiations	43
4.2 Property-instances	45
4.3 Assumptions within the Mental Causation Debate	48
4.4 Characterising Particulars and Psychophysical Reductionism	50
4.5 Property Analysis	57
4.6 Singularist and Generalist Theories of Causation	60
5 Why the Properties of Causation are Characterising Particulars	65
5.1 States of Affairs	65
5.2 Kimean Events	68
5.3 The Properties of Causation and Three Different Understandings of 'Substance'	69

5.4 Are the Properties of Causation In Re Universals or Characterising Particulars?	75
6 The Homogeneity of the Causal Relata	80
7 Summary	85

Part Two

Property Analysis

1 Property Analysis and the Mental Causation Debate	88
2 Do Characterising Particulars have Determinate Identity and Countability?	92
3 Co-instantive Supervenience and Property Dependence	98
3.1 Supervenience and the Mental Causation Debate	98
3.2 Co-instantiation Theories of Supervenience	101
3.3 Property Dependence and Co-instantiation Theories of Supervenience	103
3.4 Co-instantiation Theories of Supervenience with Supplementation	109
4 A Semantical Analysis of Characterising Particulars	113
4.1 A Semantical Analysis of Characterising Particulars and Property Layers	113
4.2 A Semantical Criterion of Characterising Particulars' Identity and Existence	114
4.3 Applications within the Mental Causation Debate	118
5 A Causal Analysis of Characterising Particulars	128
5.1 A Causal Criterion of Characterising Particulars' Identity and Existence	128
5.2 Applications within the Mental Causation Debate	130
6 Mereology	141

Part Three

Causal Closure Principles

1 Causal Closure Arguments	151
1.1 Preliminaries	151
1.2 Two Different Kinds of Causal Closure Argument	152
1.3 How to Assess the Plausibility of a Causal Closure Principle	157
1.4 Outline	160
2 A Priori Arguments for Causal Closure Principles	162
2.1 A Mark of the 'Physical'	162

2.2 A Rejection of A Priori Arguments for Causal Closure Principles	167
3 An Appeal to the Discipline of Physics	171
3.1 Sturgeon's Dilemma	171
3.2 Hempel's Dilemma	174
3.3 Melynck's Appeal to Current Physics	177
3.4 Smart's Appeal to Current Physics	179
3.5 Future Physics	183
3.6 Papineau and 'Closed Physics'	184
4 Empirical Support for 'No Effect has a Non-Physical Cause'	188
4.1 An Energy Transference Theory of Causation	188
4.2 Fair's Solution to the Problem of Hume's Similarity Circle	191
5 Empirical Support for 'No Physical Effect has a Non-Physical Cause'	195
5.1 The Conservation Laws of Physics	195
5.2 The Premise that Energy is Essentially Physical	197
5.3 The Premise of the Energy Transference Theory of Causation	199
6 Empirical Support for Weak Causal Closure Principles	202
6.1 The Problem with Hart's Theory of Psychophysical Interactionism	202
6.2 The Energy Transference Theory of Causation & Assumptions within the Mental Causation Debate	205
6.3 The Rejection of an Energy Transference Theory of Causation When Applied to Psychophysical Interactions	209
6.4 Working Towards a More Plausible Model of Causation	212
6.5 Interactive Mentalism	220
Summary	231
Bibliography	237

Notation

Properties (neutral)	P1, P2, P3...
Universals	U1, U2, U3...
Characterising particulars	CP1, CP2, CP3...
Predicates	<i>p1, p2, p3...</i>
Events	e1, e2, e3...
Event-types	E1, E2, E3...
Facts	f1, f2, f3...
Substances (metaphysically neutral)	s1, s2, s3...
Substances	S1, S2, S3...
Thin particulars	St1, St2, St3...
Times	t1, t2, t3...

Introduction

In this thesis I attempt to show that many of the fundamental issues within the mental causation debate are metaphysical ones, where metaphysics is the science of being. For this reason, to advance a plausible response to the problem of mental causation requires an extensive excursion into ontology. Attempts to respond to the problem of mental causation, which divorce themselves from ontology, or which are based upon ad hoc ontological assumptions, will inevitably prove unsatisfactory.

1. The Problem of Mental Causation

What is the problem of mental causation with which I shall be concerned? Let us start with the modest proposal that there are intentional mental states. That is, there are such things as a person desiring a glass of water, or believing that it is going to snow, or hoping that they will not be late. When we reflect upon ourselves and our relation with the world, little seems more obvious than the claim that such mental states cause physical states (such as bodily movement). Thus, for example, my belief that it is going to snow is cited amongst the causes of my action of putting on a pair of gloves when I go outside. My hoping that I will not be late is cited amongst the causes of my running for the bus. Indeed, our very idea of causation seems to stem from our experience of ourselves as causal agents. Contrary to Hume, our idea of causation does not seem to be merely derived from regular succession, but rather our experience of ourselves making a causal difference within the world. There seems to be a fundamental difference between experiencing and doing.

The premise that there is psychophysical causation is appealed to in causal closure arguments, all of which yield the conclusion that mental causes are identical with physical causes. This is a family of arguments, whose differing structures depend upon the strength of the causal closure principle to which they appeal. The most popular form of this argument is the argument from causal overdetermination.¹ It combines the premise of psychophysical causation with two further, seemingly plausible, premises:

¹This is the type of causal closure argument that shall be assumed in Part One and Part Two. In Part Three, stronger causal closure arguments will also be considered.

1. Mental causes have physical effects
2. Every physical effect has a set of physical causes which together are sufficient for its occurrence.
3. There is no systematic causal overdetermination.

Mental causes are identical with physical causes.

Given the causal closure principle that this argument appeals to, we will never need to appeal to non-physical causes in order to give a complete causal account of physical effects, because a physical effect that has a non-physical cause will also have a complete physical cause. Given the denial of systematic causal overdetermination, there is no causal work left for non-physical causes within the physical domain. Hence, to avoid the rejection of psychophysical causation, mental causes must be identical with physical causes.

It is widely acknowledged that the argument from causal overdetermination leads to problems if mental and physical properties are assumed to be instantiated by distinct substances. Hence, if one maintains the kind of substance dualism, usually (although questionably) associated with Descartes, in which a person is identified with a purely non-physical substance, then if the physical world is a causally closed system, one seems forced to conclude that mental substances are wholly epiphenomenal within the physical domain.

But the problem is not confined to such a substance dualism. Importantly, it is also a problem for a substance monist, who maintains a property dualism. Even if mental and physical properties belong to the same substantial particular, that is, even if human beings are just complex collections of purely physical parts, the problem is not removed. This is because, although there is much disagreement about what the causal relata are, most would agree that properties play a central role within causation: the causal relations that an entity can enter into depends upon what that entity is like, and hence which properties characterise it. Consequently, to identify mental substances with physical substances is not to remove, but merely to relocate the problem, for unless mental and physical properties are also identified, questions about the causal redundancy of the mental reappear at the level of properties. One must consider whether the mental properties of a substance are required to play any causal role within the physical domain. If they are, then given the denial of systematic causal overdetermination, this is to reject the causal closure

principle. Alternatively, if mental properties are causally redundant within the physical domain, the resulting property epiphenomenalism is no less serious than the original substance epiphenomenalism that substance dualism is charged with.

We are now in a position to map out the various positions within the mental causation debate. To a certain degree these can be distinguished by their differing responses to the argument from causal overdetermination. Let us first clarify our distinction between substance dualism and substance monism. Descartes' substance dualism is not the only form of substance dualism. Substance dualists differ in their understanding of what it is to be a person. While Descartes identifies a person with a purely non-physical substance, according to a second view, a person is a combination of a physical body and a non-physical soul. However, to be a substance dualist one need not be committed to the existence of a purely non-physical substance. Lowe (1996, 2000b), for example, maintains that persons are simple substances possessing both physical and mental properties. Although not all substance dualists are committed to the claim that mental and physical properties are housed in distinct substances, all agree that mental properties are not ways that a body is, nor ways that any part of it is. As the body is composed of matter, all forms of substance dualism therefore conflict with the principle of universal physical composition (UPC), where UPC is the principle that all objects are wholly composed of matter.

Substance monists accept UPC.² Given UPC, those substances to which mental properties belong are really just arrangements of matter. UPC is commonly combined with a multileveled model of the world in which all objects form a single hierarchical structure generated by the mereological relation, the objects at each level being complex structures of those at the level below it, and the lowest level consisting of the smallest parts of matter. At each level, new properties that are 'characteristic' of that level are encountered. For example, at the sub-atomic level we find properties whose values are determined by a system's quantum mechanical state, such as position, spin, energy, and number. Electrical conductivity, viscosity, and density are found at the molecular level. Mental properties are found at the level of higher organisms. The study of each level is thought to

² In some respects UPC might be thought to be too strong to characterise substance monism. The substance monist will want to admit that there are many kinds of things, such as irrational numbers and the rules of cricket that do not appear to be composed of anything at all. But the alternative weaker position, that all particulars *that have parts* must have only physical parts, although excluding the possibility that the self may be composed of some immaterial stuff, is compatible with those forms of substance dualism that do not wish to maintain that the self is composed of anything at all. At least for the purpose of the philosophy of mind, UPC is more appropriate.

be the task of a particular domain of science, its aim being to identify the laws that govern the behaviour of objects at the relevant level. As the discipline that traditionally studies the basic parts of matter out of which all material things are composed is physics, UPC suggests that all objects either are, or are composed of, the objects described by physics. It is, however, important to recognise that UCP need not be combined with a layered model, in which there are, not only basic particles, but also the aggregates of these particles, and the aggregates of these aggregates. There may only be a single level of objects, properties and relations which consists in the most basic particles, that is, those of physics. Given such a view, mental properties and relations are identified with neurological properties and relations, and these are ultimately identified with those of physics.

With this distinction we can now consider what I shall refer to as 'mentalism'. A mentalist may maintain either a substance pluralism or a substance monism. According to the mentalist, either mental properties do not belong to the body, or although mental properties are properties of the body, they are not merely distinct from physical properties of the body, but also ontologically independent of them.³ The term 'anti-physicalism' shall be used to refer to the more general thesis of which the former is a species.

Mentalism can be interpreted in a number of different ways, which differ in the causal status allowed to mental properties. A mentalist may maintain what Seager (1991, p.16) refers to as an 'empty dualism': mental properties do not causally affect the physical domain and either an epiphenomenalism or parallelism holds. This kind of mentalism admits the premise of causal closure, whilst denying the premise of psychophysical causation. Given the high plausibility of the claim that there is psychophysical causation, this is not a position that the mentalist should willingly accept.

Alternatively, and far more plausibly, one may hold an interactive mentalism. An interactive mentalism may be combined with either a substance pluralism or a substance monism. According to the first position, although mental properties are not properties of the body they make a causal difference to the body. According to the second position, certain configurations of physical properties give rise to mental properties that make a real causal difference. It is not the case that mental properties and relations do not exist in as metaphysically robust a sense as the physical properties and relations belonging to the

³ Note, that one can be a mentalist without being a dualist, for the mentalist's claim is that there are at least two domains of ontologically independent phenomena.

parts from which they are composed. In particular, even though a whole depends for its existence on the parts that compose it, this is not to suggest that the properties of a whole cannot have independent causal powers that exist over and above the causal powers of the properties of its parts and their relations. The interactive property mentalist will maintain that mental properties have independent causal powers, which can affect lower levels of properties on the macro-micro hierarchy of objects. That is, mental causes exert 'downwards causation.' Hence, what it is that unites interactive substance mentalism and interactive property mentalism, is the claim that mental states have full-blooded, independent causal effects within the physical domain, these effects being neither grounded in nor determined by the causal powers of physical properties.

How can the interactive mentalist respond to the argument from causal overdetermination? He has a number of different options. In the first place, he may deny one of the premises of the argument. That is, he may deny either the premise of causal closure, or the premise of the denial of systematic causal overdetermination. Alternatively, although accepting the premises of the argument, the interactive mentalist may deny that the conclusion that mental causes are identical with physical causes follows from these premises. The premises are together consistent with mental causes having independent effects within the physical domain. This may be based upon an appeal to the heterogeneity of the causal relata or the heterogeneity of the causal relation.

Now the problem of mental causation is not just a problem for the mentalist. The minimal requirement within any physicalism is that:

P: All properties and relations are in some sense physical properties and relations. (Note, that if all properties and relations are physical, it follows that all objects are physical).

Eliminativism, psychophysical reductionism, and non-reductive physicalism provide distinct articulations of P, differences between them hinging on their attitude to mental properties and relations. The first maintains that mental properties and relations should be eliminated. The second maintains that they should be identified with physical properties and relations. The third advances a property dualism, but arguably differs from property mentalism in its suggestion that mental properties and relations are in some way dependent upon physical properties and relations. Most importantly, it differs from an interactive property mentalism, because the non-reductive physicalist maintains that the

causal powers of a mental property are entirely determined by or grounded in the causal powers of physical properties.

Physicalism is not entailed by the causal closure principle: epiphenomenalism, parallelism and some forms of interactive mentalism also accept the causal closure principle. But, for the purpose of this discussion, we can assume that all physicalists are committed to the causal closure principle, although note that it is possible to be a 'physicalist', but reject the causal closure principle, given that the term 'physical' within the causal closure principle should be interpreted in a narrow sense. One's physicalism largely depends upon what it is that one considers the term 'physical' to include. There are two broad options. The first advances a narrow interpretation of the term 'physical' which includes only the set of properties and relations appealed to within the discipline of physics. The alternative, broader interpretation of the 'physical' also includes those properties and relations within the higher-level sciences. A physicalist might adopt a broad understanding of the term 'physical', and maintain that there are emergent physical properties. However, my aim is not to critically address the entire spectrum of physicalist positions. For the purpose of this debate, I am only concerned with the group of physicalist theories that maintain that the only physical causes that have independent causal powers are those within physics. That is, those physicalists that maintain a causal monism within the physical domain. Hence, we can assume that all physicalists are committed to the causal closure principle. Certainly, it will be argued that it is hard to see how one could motivate a physicalism which did not assume the causal closure principle, as the causal closure of the physical domain is the central premise within the best arguments for physicalism.

Thus all physicalists are committed to the causal closure principle. Why then, do they face a problem with the argument from causal overdetermination? Whilst eliminativism rejects the premise of psychophysical causation — there are no mental entities, and hence there is no mental causation — all other forms of physicalism desire to preserve mental causation. Hence, along with their acceptance of the causal closure principle, most physicalists also accept the premise of psychophysical causation. Due to their desire to preserve psychophysical causation, whilst staying true to physicalism and hence the causal closure principle, the obvious thing for the physicalist to do, would be to go along with the conclusion that the argument from causal overdetermination leads him to, and hence maintain a psychophysical reductionism. The causal closure principle is true, and yet mental properties can make a causal difference within the physical domain, because mental properties are identical with physical properties.

But psychophysical reductionism faces its own serious problems, in particular, that of the argument from multiple realisability. Many philosophers of mind have argued that mental properties are realisable by different physical properties. Though the property of being in pain, for example, might be realised in humans by one physical property, pain might be realised by a different physical property in animals, and still yet another in extra-terrestrial beings. This point generalises to all mental properties. The multiple realisability of mental properties by physical properties leads to the rejection of their identification. Mental and physical properties must be distinct.

The difficulties that confront the psychophysical reductionist reopen the problem of mental causation for the physicalist. Rather than psychophysical reductionism, most physicalists adopt a non-reductive physicalism. Mental and physical properties are distinct properties of the body. However, non-reductive physicalism is not a property mentalism. It differs from interactive property mentalism, because contrary to the interactive mentalist, the causal powers of a mental property are not independent of the causal powers of physical properties. It differs from epiphenomenalism and parallelism, in its commitment to the premise of psychophysical causation. And it arguably fundamentally differs from all forms of mentalism, in its claim that the physical determines everything. Hence, the non-reductive physicalist hopes to deny the identity of mental and physical properties, whilst staying true to the primacy of the physical domain, and also maintaining the premise of psychophysical causation.

But as with all other forms of property dualism, the non-reductive physicalist faces the problem of mental causation. If mental properties are not identical with physical properties, but the physical domain is causally closed and there is no systematic causal overdetermination, how can there be psychophysical causation? That is, if mental properties are distinct from physical properties, how can one accommodate the conviction that the physical world is causally closed and also the common-sense view that there is psychophysical causation? No matter how intimately related mental properties are to physical properties, the fact remains that mental properties are distinct from physical properties. Hence how is the non-reductive physicalist to avoid property epiphenomenalism if he is to stay true to the premise of causal closure?

The non-reductive physicalist must respond by showing that one can accept the premises of the argument from causal overdetermination, whilst denying its conclusion. Although mental properties and hence mental causes are distinct from the physical, the causal status of a mental cause is in some sense nothing over and above that of a physical

cause. But can the non-reductive physicalist really forge a route between property epiphenomenalism and a 'full-blooded' interactive property mentalism, despite denying the identity of mental and physical properties? The argument from causal overdetermination provides a problem not only for the mentalist, but also, given the unacceptability of identifying mental and physical properties, the physicalist.⁴

2. Metaphysics and the Problem of Mental Causation

In this discussion, I hope to show, that to be plausible, a theory of mental causation must be developed from within a sound metaphysics. Those in the mental causation debate have tended to keep their distance from metaphysical issues, but this can only lead to error and confusion, for the fundamental questions in the mental causation debate are, I would suggest, metaphysical ones. Hence, it is only with metaphysical clarity, that any clarity can be gained in the mental causation debate. Many of the problems that exercise those within the mental causation debate arise from a failure to base their arguments in a sound metaphysics. And whilst sloppy metaphysics has led to a plethora of unpersuasive positions, grounding the mental causation debate in a sound metaphysics leads us to see what the mental causation debate is really a debate about, and who it is most plausibly a debate between. It also leads to the clarification and questioning of certain premises in the argument from causal overdetermination, hence allowing us to see the form that a solution to the problem of mental causation should take.

What do I understand metaphysics to be? With Lowe (1998, 2001b), Armstrong (1997), and Martin (1993), I understand metaphysics' concern to be with ontology, with the fundamental structures of reality. This explanation of metaphysics admittedly lacks detail. As Lowe argues (1998, ch. 1, p. 2), to attempt to offer a definition of 'metaphysics' or specify its subject-matter by listing its topics, would invariably be to rule out certain metaphysical positions and ignore others.

One can, however, go some way to giving a general explanation of what it is to offer a metaphysical system. Ontological categories are different categories of being. Whether an

⁴ Note that there are other problems of mental causation, in addition to the problem of causal overdetermination. In particular, there is the problem of the causal relevance of mental content. That is, the problem of how the contents of propositional attitude states can be causally relevant to the effects that those states have. For a discussion of this issue, see, for example, Lowe (2000b, ch. 4).

entity belongs to a certain ontological category is determinable a priori by consideration of its existence and identity conditions. (See Lowe, (forthcoming (a), p.1)). Hence, for example, whether one should admit the ontological category of properties, and indeed what properties are, is to be determined by considering whether the ontological category of properties could exist, and thus their existence and identity conditions.

To offer a metaphysical system is to offer an ontological system. Different ontological systems will appeal to different ontological categories. For example, whilst Armstrong's (1997) ontological system appeals to states of affairs as ontologically additional entities, Campbell (1991)'s trope ontology does not. They may also understand some ontological categories within this system to be more basic than others. For example, Armstrong understands the ontological category of states of affairs to be more basic than the ontological categories of substance and property to which his system also appeals. Whether one should admit an ontological category within one's ontological system is to be determined by considering whether such an ontological category could play a non-redundant role in one's ontological system, and whether it could co-exist with the other ontological categories that one includes within this ontological system.

Consideration of the ontological systems that could exist is a purely a priori pursuit. The role of empirical science is then to determine the kind of things that do exist in the actual world. But before we can decide what actually does exist, we must determine what could exist, because as Lowe explains 'empirical evidence can only be evidence for the existence of something whose existence is antecedently possible.' (Lowe, 2001b, pp. 3-4). Consequently, in order to address the problem of mental causation one cannot turn one's back on metaphysical issues and instead turn to the empirical sciences, because metaphysics provides the very ontological framework in which to locate empirical truths. Before one can appeal to empirical considerations in order to advance a particular theory of mental causation, one must first have a clear understanding of the underlying ontology.

Although the specific reasons why the mental causation debate should be concerned with metaphysics are many and diverse, one can, I think, divide the metaphysical issues that are of importance to the mental causation debate into three broad groups. The first centres upon the question of what the causal relata are. What causation is a relation between, depends upon the ontological system within which one is basing causation, and hence the ontological categories to which one appeals. The mental causation debate cannot ignore the issue of the causal relata, because the interpretation of the argument from causal overdetermination, and the possible responses to it, depend upon it.

The second area of importance to the mental causation debate is property analysis. That is, a consideration of the existence conditions for properties, the identity conditions for properties and the further kinds of relations, for example, dependence relations, that properties can enter into with each other and other ontological categories. Given that it is properties that play the central role in causation, that is, it is a property that makes a particular cause the kind of cause that it is, then it follows that what it is for a cause (or effect) to exist (partly) depends upon what it is for a property to exist, and what it is for two causes (or effects) to be identical with one another (partly) depends upon what it is for one property to be identical with another. Hence, what it is for a mental cause to exist, and what it is for a mental cause to be identical with a physical cause, depends upon one's analysis of properties. Property analysis will also allow us to determine whether one property can depend upon another, and if so, in what sense; hence allowing us to address the question of whether mental properties depend upon physical properties and what this would entail for the mental causation debate.

Finally, I would suggest that the mental causation debate is greatly affected by what one understands the causal relation to be. Whether a closure principle that is strong enough to allow one to advance physicalism, can plausibly be advanced, heavily depends upon the theory of causation in which one is embedding one's discussion of psychophysical causation.

Clearly, it would be incorrect to tailor one's metaphysics to suit some position within the mental causation debate. A theory of mental causation that is based upon ad hoc metaphysical assumptions will prove unsatisfying. Consequently, one's theory of mental causation must be embedded within a metaphysical system that can be independently motivated. Discussion of these three metaphysical areas demands a certain order of exposition. How to analyse properties, depends upon what properties are, which is partly determined by one's theory of causal relata. One's theory of causation is also greatly affected by one's understanding of a property. Consequently, from a metaphysical perspective, this discussion should begin with a discussion of the causal relata, then move on to property analysis, and then finish with a discussion of the causal relation. Fortunately, this fits in well with the order in which matters need to be raised within the mental causation debate. With a plausible analysis of the causal relata, one can go on to provide a clear formulation of the argument from causal overdetermination. One can also rule out certain positions within the mental causation debate that are based within an implausible theory of the causal relata. A plausible analysis of properties leads us to see

that the true contenders within the mental causation debate are interactive mentalism on the one side, and psychophysical reductionism on the other. Finally, consideration of the causal relation helps to reveal that it is far from evident that a causal closure principle that is strong enough to support physicalism is plausible, hence, suggesting that one should in fact advance an interactive mentalism.

Part One

THE CAUSAL RELATA

Mental Causation and the Causal Relata

What are the causal relata? Those within the mental causation debate cannot divorce themselves from this question, and neither can they hope to provide a theory of mental causation that is somehow neutral between the various candidates for the causal relata. This is because what the problem of mental causation is actually a problem about, and the possible ways of responding to it, depend upon what causation is a relation between; one's theory of the causal relata provides the very framework for one's theory of mental causation. A belief to the contrary has led, not only to poorly formulated theories of mental causation, but the misrepresentation of those positions within the mental causation debate that are ontologically serious. On the other hand, one obviously cannot choose a theory of the causal relata simply because it promises to offer a solution to the problem of mental causation. One's theory of the causal relata must be independently plausible.

The aim of Part One of this thesis is two-fold. I hope to show how certain assumptions within the mental causation debate about the causal relata have influenced the interpretation of the problem of causal overdetermination and the responses offered to it. In conjunction with this, I shall defend a theory of the causal relata that rejects some of these assumptions, and consider how this alternative picture of the causal relata affects the mental causation debate.

After raising some general considerations about the causal relata in Section 1, I go on to defend three claims concerning the causal relata. Firstly, the causal relata are entities that have properties and these play the central role within causation; it is properties that make the causal difference within a causal relation (§2). Secondly, causes do not have epiphenomenal properties — the causal relata are 'fine-grained' (§3). Thirdly, the properties that the causal relata instantiate — the properties of causation — are particulars, not universals (§4 and §5).

The first claim is widely accepted both within discussions of the causal relata and the mental causation debate. One notable exception is Davidson (1980d, 1980e and 1980f, 1993), whose solution to the mental causation debate — that of anomalous monism — is

in fact dependent upon his rejection of this claim. Davidson's anomalous monism and the debate that surrounds it, serve to provide an excellent demonstration of the fact that a failure to take into account the ontological system within which a theory of mental causation is based can lead to totally inappropriate criticisms of it. Despite the fact that Davidson's theory of the causal relata is crucial to his response to the problem of mental causation, it is commonly overlooked within discussions of anomalous monism. Anomalous monism is accused of entailing property epiphenomenalism, but given Davidson's understanding of the causal relata such accusations are inappropriate.¹ Davidson's anomalous monism should in fact be rejected because of the implausibility of the ontological system within which it is based.

My claim that the causal relata are fine-grained follows from considerations of the causal specificity of causation together with its transitivity. The causal relata are what I refer to as 'property-instantiations'. This serves to simplify the mental causation debate somewhat, for it yields a straightforward interpretation of the argument from causal overdetermination. The question simply becomes the question of how mental causes, if not identical with physical causes, can have physical effects. If mental causes are identical with physical causes, there can be no further question of whether the 'mentalness' of the mental cause is in some sense causally redundant.

Finally, I argue that a plausible understanding of property-instantiations as the causal relata, requires that the properties of causation are particulars. Despite the fact that the nature of the properties of causation is rarely discussed within the mental causation debate, the implicit assumption is that they are universals. I argue that, contrary to Robb and Heil (Robb (1997), Robb (2001), Heil and Robb (forthcoming)), the shift from universals to particulars does not make it more plausible to identify mental causes with physical causes and hence advance a psychophysical reductionism. However, it does have a number of very important and wide-ranging effects upon the mental causation debate. In the first place, one's analysis of properties is crucial to the mental causation debate, but how to analyse properties depends upon what properties are. Secondly, if the properties of causation are particulars, this affects one's analysis of the causal relation itself, which in turn affects one's understanding of what it is for there to be psychophysical causation, and hence, I would suggest, the plausibility of the premise within causal closure arguments that the physical domain is causally closed.

¹ I am indebted to Crane for initiating this line of thought. See Crane (1992, p 188-190) & Crane (1995, p. 226-229).

The first part of my thesis ends with a discussion of E. J. Lowe's (1989, 2000) theory of interactive mentalism. Lowe responds to the problem of causal overdetermination by accepting its premises, but rejecting its physicalist conclusion. He does this by defending a psychophysical causal interactionism based upon the denial of the homogeneity of the causal relata. Mental and physical causation are, according to this account, causation by different kinds of entity. Lowe's discussion reveals yet another way in which the causal relata may play an important role within the mental causation debate. However, the distinction between event causation and fact causation that Lowe's theory of psychophysical interactionism requires, is, I argue, not one that can be ontologically defended.

1.1 Preliminaries

To avoid potential misunderstanding, I shall first make clear three of my fundamental assumptions about causation. I assume that causation is a relation. This is denied by Achinstein (1983) who appeals to the phenomena of emphasis within causal claims to show that singular causal sentences are non-relational. I shall not criticise Achinstein's argument here, but refer the reader to other sources (For example, see Hausman (1998 pp.23-5)).

Secondly, causation is an objective relation. Most causal relations are mind-independent. For example, without cognisers to think about it, the wind would still cause twigs to snap, for events such as snappings and substances such as twigs exist independently of minds, as does the causal relation that holds between them. Furthermore, through metaphysical enquiry one can find out about the causal relation as it is in itself. Hence, I consider idealism and phenomenalism which maintain that nothing is mind-independent to be false, and also the Kantian view of metaphysics which maintains that metaphysics can only reveal our thoughts about objective reality, rather than anything about how objective reality really is.

Thirdly, as causation is an objective relation, what it is a relation between will be determined by what there is in the world. A discussion of the causal relata cannot be conducted in isolation from ontological considerations. Rather, an account of the causal relata should begin with the setting out of one's ontology and be developed alongside it.

What one considers the causal relata to be will depend upon what one considers to be ontologically basic and dissatisfaction with an account of the causal relata will in many cases stem from a dissatisfaction with the ontology from within which it is developed.

1.2 The Particularity of the Causal Relata

A number of ontological categories are clearly unsuited to the role of causal relata. Singular causation is a relation between entities existing at the level of particulars, where the notion of a 'particular' is to be contrasted with that of a 'kind'. Hence, neither universals (e.g. redness, heat, etc.) nor substantial kinds (e.g. apples, cats, etc.) are the causal relata. Equally, if events are the causal relata, they must be event tokens rather than event types. Therefore, if events are literally repeatable, as Chisholm (1970) maintains, they cannot be the causal relata, although particular instantiations of such events could be.

Arguably, singular causation must also be a relation between concrete entities, where to be concrete is to possess a spatio-temporal location. Indeed, this may be thought to follow from the fact that the causal relata are particulars, as according to one understanding of a 'particular' to be a particular is to possess a single location in space at a time. But this understanding of a particular is questionable. Although it is true that, in so far as a particular has a spatial determination, it must possess an individual location in space at a time, one can not rule out a priori the possibility of there being particulars that possess an individual location in some higher dimensional space, or indeed the possibility of there being particulars which are not located in any space, e.g. Cartesian substances, entelechies, etc. Hence to be a particular cannot be to have a single location in space at a time, or at the very least this understanding of a particular cannot derive from a conceptual consideration of particularity.

For similar reasons, the claim that causation must be a relation between spatial entities is too strong. It is true that the causal relata must occur in time. This is not merely because generalist theories of causation incorporate the singularist requirement of temporal locatedness, appealing to temporal priority in order to explain causal priority. Rather, it is motivated by the more general consideration that causation is intimately bound up with the notion of change, which is itself intimately connected to the notion of time. However, it is

not a part of the concept of causation that cause and effect must be spatially located. It is true that some theories of causation maintain that the causal relata are spatially located. For example, in the *Treatise* Hume claims that cause and effect must be spatially contiguous, and hence located in space (Hume (1738, Book I, Part IV, § 5)). Nor is this requirement unique to the naïve regularity theory. The spatial contiguity of cause and effect is also, for example, a consequence of Ducasse's (1968) reductive singularist theory of causation. More generally, without the requirement of spatio-temporal contiguity, questions concerning which is the real cause in cases of pre-emption become harder to deal with for some theories of causation.² But (contrary to Hume) this cannot be an a priori constraint. As a matter of fact all entities that enter into causal relations may be spatially located, but this is to make a very different point. One cannot rule out a priori the possibility that some entities that enter into causal relations lack a location in this space, or indeed a location in any space at all.

What one can say is that if there are causal relations between non-concrete entities, their non-concreteness must not derive from the nature of the causal relata itself, but the non-concreteness of the object that in this particular case the causal relation is between. If in fact all objects are concrete objects, then all causal relations must be between concrete entities. To this extent events are plausible candidates for the causal relata. This is because events have a location in space in virtue of the spatial locatedness of the objects which they are changes (or unchanges) within. Hence, for example, the event of the bridge collapsing is itself concrete, because the bridge is a concrete object. Unlike an event, if, for example, a set has wholly concrete objects as its members, despite this fact the set is still itself an abstract entity. For this reason, causation is not a relation between sets, or indeed any other entity that is abstract in this sense. One can therefore dismiss as possible candidates for the causal relata Taylor's (1985) theory of events (which identifies events with sets) and those theories of facts that consider facts to be abstract unities or to be composed of purely abstract parts.

1.3 Causation as a Relation between Events

Causation is a relation between particulars, but what sort of particular? The most popular claim is that causation is a relation between 'events', where events are particulars (contra Chisholm) and which when involving concrete objects are themselves located in space

² See Hausman (1998, p.267)

(contra Taylor). However, there is little further agreement about what events are. Of those philosophers who admit the ontological category of properties, most would agree, that, as with objects, the properties that a definite description of an event refers to do not exhaust the properties of that event. Hence, 'Kate's wedding' is an event that has, not only the property of being a wedding, but also occurring on a Saturday, of having two hundred guests, etc. In this respect, events are like objects, but differ from states of affairs, which are built out of, for example, a substance and a property that this substance exemplifies. However, not all theories of events are thick-grained. In particular, Kim's (1993b, 1993c) theory of events comes close to suggesting that an event is a substance exemplifying a single property, and in doing so arguably obliterates the distinction between an event and a state of affairs.³

Another point of disagreement regarding events is whether there is a distinction between an event and a state. Events are commonly thought to differ from objects because the temporal boundaries of events contribute to their identity conditions, whilst only the spatial boundaries of ordinary concrete objects contribute to their identity conditions. Although objects have a location in time, one object cannot be a temporal part of another; an object is wholly present at any instant of time at which it exists. Events, on the other hand, have temporal parts and are never wholly present at any one instant. Thus, for example, if the wilting of a flower takes place between t_1 and t_2 , the flower is wholly present at every instant within this time interval. But the event, the wilting of the flower, is never wholly present at any instant between t_1 and t_2 . Hence one might say that an event involving a substance is a change in its properties, that is the acquiring or gaining of a property by a substance, and a state is an unchange in its properties, that is the having of a property by a substance.⁴ In this sense, a mental episode such as feeling a pain is an event, for it takes place over time. On the other hand, having a belief is a mental state for it clearly does not have temporal parts. But not all philosophers identify events with changes in substances. Some collapse the distinction between event and states, suggesting that an event can be wholly present at a single instant in time.⁵

³ I oppose this last point — this difference between an event and a state of affairs is not crucial. See Part 1, §5.

⁴ Lombard (1986) maintains that events necessarily involve change. Indeed, he considers that one cannot get a grip on the concept of an event without seeing it as bound up in the concept of change.

⁵ For example, Kimean events are exemplifications of properties by substances at a time. See Kim (1993b, 1993c)

But I would suggest that by far the most fundamental disagreement regarding events concerns the ontological categories to which they are related, and the ontological status that they should be given. From the claim that events are the causal relata one can conclude little unless one has first examined the ontology within which the relevant theory of events is based. To give but a few examples of the different ontological understandings of an event, Campbell (1991) and Williams (1966) maintain a bundle of tropes theory in which events are identified with lively tropes (Williams), or trope sequences (Campbell). Within this one-category ontology one might describe the causal relata as events, but events are not ontologically additional to tropes and it is tropes that are ontologically basic.

Alternatively, Kim (1993b, 1993c) maintains a two-category ontology and considers events to be the causal relata, where an event exists if a substance exemplifies a property at a time. For Kim, events are ontologically additional to the categories of substance and property. Armstrong (1997) also adopts a two-category ontology, but unlike Kim identifies a property-exemplification not with an event but a state of affairs. Preserving the connection between events and changes, Armstrong maintains that an event is a 'succession of states of affairs.' (1997, p.206). For Armstrong, although states of affairs are ontologically basic entities, events are not. Finally, given a two-category ontology, one may reject Armstrong's claim that property-exemplifications are states of affairs, and Kim's claim that they are events, and instead identify them with states of objects. Preserving the distinction between a state and an event, an event is a complex of property-exemplifications occurring within a spatio-temporal zone.

Lastly, events may be incorporated into a nominalist ontology. There are no properties; there are just substances. Hence, for example, according to Quine (1960, p.171), an event is the content of some portion of space-time. But given Quine's nominalist ontology this is not to be identified with the previous position that an event is the complex of property-exemplifications occurring within a spatio-temporal zone. To which ontological categories events are related, and in particular whether or not events are related to the ontological category of properties, is the central question that lies behind our discussion of Davidson's anomalous monism.

2

Davidson's Anomalous Monism and Property Epiphenomenalism

2.1 Anomalous Monism

Let us assume that events are the causal relata. Given an event monism, mental events will be causally efficacious because they are token identical with physical events. But if events are causes in virtue of the properties that they instantiate, unless mental properties and physical properties are also identified, questions about the causal redundancy of the mental reappear at the level of properties. One must consider whether mental events are ever causes of physical events in virtue of their mental properties. If the mental properties of a mental event do make a causal difference, then, unless one admits systematic causal overdetermination, this is to violate the causal closure principle, for according to it an event's physical properties are sufficient for the causal effects that that event has within the physical domain. Alternatively, if the mental properties of an event make no causal difference to the physical effects that the event has, then mental properties have the status of epiphenomena. Consequently, to identify mental events with physical events whilst distinguishing mental properties from physical properties is not to remove but to merely relocate the problem of mental causation. For this reason, the non-reductive physicalist who identifies token mental events with physical events, but maintains a type dualism, can plausibly be accused of property epiphenomenalism.

Davidson's anomalous monism (1980d, e and f, 1993) is thought to be the classical example of this form of non-reductive physicalism. It advances the reduction of mental events to physical events, whilst rejecting the reduction of mental concepts to physical concepts, and is derived from three premises.

- (1) The Principle of Causal Interaction (CI): At least some mental events are causally related to physical events.
- (2) The Principle of the Nomological Character of Causality (NCC): All singular causal relations are backed by strict laws.

(3) The Principle of the Anomalism of the Mental (AM): There are no strict psychophysical laws.

According to Davidson, the basic causal relata are events. Given NCC, two events are causally connected if and only if they are connected by a strict law. However, the strict law need not be formulated in the same terms as the causal claim. Rather, causally related events must have descriptions under which they instantiate a strict law. Given CI, mental and physical events must therefore be describable in terms that allow them to fall under a strict law. AM entails that a psychophysical causal claim and the relevant covering law cannot be formulated in the same terms; the latter must be formulated in purely physical terms. Consequently, any mental event that causes a physical event must be characterisable in physical terms and therefore be physical. Hence, mental events are physical events. On the other hand, as there are no strict psychophysical laws that would support the reduction of mental concepts to physical concepts, anomalous monism leads to the rejection of any conceptual reduction.

It is widely assumed that anomalous monism leads to epiphenomenalism.⁶ Contra Davidson (1993, p. 3), the criticism is not that the three premises of anomalous monism are mutually inconsistent, that is that NCC and AM are together inconsistent with CI and lead to an event epiphenomenalism. According to his critics Davidson 'can (and does) deny token epiphenomenalism' (McLaughlin (1993, p. 28)) for they consider NCC and AM to be consistent with the token identification of mental events with physical events. Rather, their charge is that of property epiphenomenalism. Given anomalous monism, how can the mental properties of an event make a causal difference in the physical domain; for how can anomalous properties of an event be causal properties, given NCC? NCC and AM together entail that mental events are causes only as they instantiate physical laws and therefore that mental events are only causes in virtue of the properties picked out by physical descriptions. It is therefore only in virtue of their physical properties that mental events cause physical events.

As is customary, I have used the locution 'in virtue of' to express the criticism of property epiphenomenalism. There are, however, two different interpretations of the 'in virtue of' claim, which result in two different forms of property epiphenomenalism. Each assumes a different understanding of the term 'property'. The first form of property epiphenomenalism

⁶ For instance, see B. P. McLaughlin (1993), J. Kim (1993h), T. Honderich (1982), E. Sosa (1993) & D. Robb (1997)

understands the term 'property' in an ontological sense, whilst the second form, which I shall refer to as 'predicate epiphenomenalism', understands the term 'property' in a linguistic sense. Almost invariably the term 'property' is left ambiguous within discussions of property epiphenomenalism, so these forms of property epiphenomenalism are rarely distinguished. Indeed, discussion often unconsciously passes from one to the other, probably because it is being implicitly assumed that there is a greater correlation between predicates and properties than there in fact is.⁷

I do not consider that anomalous monism can be plausibly accused of either kind of epiphenomenalism. The first cannot legitimately be applied to anomalous monism, as it is incompatible with Davidson's ontology.⁸ And accusations of predicate epiphenomenalism, although consistent with Davidson's ontology, are ungrounded within Davidson's anomalous monism. There is, therefore, no plausible form of epiphenomenalism which anomalous monism can be accused of entailing. Philosophers of mind have mislocated the problem with Davidson's anomalous monism, which in fact lies with the implausible theory of the causal relation upon which it rests.

2.2 Property Epiphenomenalism

In 'Can Supervenience Save Anomalous Monism?' Kim argues that anomalous monism has serious problems establishing the causal efficacy of mental properties (1993h, pp.20-1). Furthermore, according to Kim, this is a concern which critics voice with an 'impressive if unsurprising unanimity' (p. 20) and indeed one which he considers Davidson to attempt to resolve, unsuccessfully, in 'Thinking Causes' by supplementing anomalous monism 'with supervenience...and perhaps also with 'non-strict laws'.' (p. 21). This criticism, although so often voiced, is misdirected. But before I consider the reasons why, I shall consider a more general issue that Kim's criticism raises.

According to Kim, given anomalous monism, mental properties lack causal efficacy and it is precisely the 'causal efficacy of properties of events' that is at issue. (p.21) (*My*

⁷ The relation between predicates and properties, and its bearing on the mental causation debate will be discussed in Part 2.

⁸ I am indebted to Crane for initiating this line of thought. Crane (1992, p 188-190) & Crane (1995, p. 226-229) is one of the few philosophers to argue that accusations of this kind of property epiphenomenalism are illegitimate when applied to anomalous monism, because they are incompatible with Davidson's ontology.

Emphasis). The property epiphenomenalism that Kim is accusing Davidson of would therefore appear to be based upon a certain understanding of the 'in virtue of' principle:

V1: If e_1 causes e_2 in virtue of having property P_1 , then P_1 is causally efficacious.

In 'Nonreductive Materialism and Mental Causation', Marras (1994) dismisses Kim's kind of criticism, not because of any particular concern about the legitimacy of V1 within a Davidsonian schema, but because of the more general worry that the form of epiphenomenalism that V1 leads to is inapplicable within any non-reductionist token identity theory. This is because: 'Properties (or types) are abstract entities, and abstract entities don't cause changes in the spatio-temporal world; only concrete *exemplifications* of properties do. Singular causal relations...occur only between tokens, individual datable events.... So the charge that anomalous monism is committed to the view that 'mental events *qua mental event* types (sic) are causally inert' simply invites the response: 'So what?' (p. 470) These considerations lead Marras to conclude that what is at issue for critics of such a non-reductive physicalism is 'not so much the causal *efficacy* of mental properties as rather the *relevance* of mental properties to the formulation of causal explanations.' (p. 473)

Of course, Marras is here assuming that properties are universals. One can ascribe properties causal efficacy if properties are particulars.⁹ But even if properties are types, Marras' conclusion is too hasty. If properties are types, then certainly, contra Kim, what is at issue is not the 'causal efficacy of mental properties'. However, just because types are not themselves causally efficacious, this does not mean that they do not play the central role within causation. And they do, provided that one maintains a generalist theory of causation.

Generalist and singularist theories of causation disagree over whether singular causation is grounded by a type level relation. The generalist maintains that it is, holding that for particular events to be causally related at the token level they must instantiate types of events that bear suitable objective relations to one another.¹⁰ For example, given a nomological theory of causation, events are causally related in virtue of exemplifying

⁹ For an understanding of tropes as property-instances see Part 1, § 4.2.

¹⁰ This distinction between singularist and generalist theories of causation is discussed in Part 1, § 4.6, where attention is also drawn to a narrower way of drawing the distinction. The above way of distinguishing singularism and generalism is advanced by Ehring (1997). See, for example Ehring (1997, p. 18).

property-types that are nomologically related. Nomological relations between property-types ontologically ground the causal relations between events. Hence, if properties are universals, although they are not causally efficacious, given a generalist theory of causation they are the things that confer causal efficacy upon events. Contrary to Marras' suggestion, it is not merely that we appeal to a nomological relation existing between universals in order to explain why a causal relation obtains. Rather, given a nomological theory of causation, for two events to be causally related is for two universals to be nomologically related.

If properties are universals, this leads to an amended version of the 'in virtue of' principle:

V2: If e1 causes e2 in virtue of e1 exemplifying universal U1 and e2 exemplifying universal U2, then U1 and U2 participate in a relationship that grounds this causal relation between e1 and e2.

I shall refer to a universal as 'causal' if it participates in a relationship that grounds a causal relation. As not all of the universals exemplified by the cause-event in a causal relation need participate in the relevant relation, this generates an epiphenomenalism, which I shall refer to as type-epiphenomenalism. If the mental is epiphenomenal in this sense this means that mental types do not ontologically ground psychophysical causal relations — the fact that an event falls under a mental type makes no difference to the causal relations that it can enter into within the physical domain. V2-epiphenomenalism is the form of epiphenomenalism that Davidson is commonly accused of. It is what McLaughlin (1993, p. 28) is referring to in accusing Davidson of 'type epiphenomenalism', what Honderich's (1982, p.62) 'Principle of the Nomological Character of Causally-Relevant Properties' leads to, and probably what Kim actually intends.

Clearly, whether an event identity theory can be accused of type epiphenomenalism depends upon the theory of causation in which it is embedded. If a token monism/ type dualism is combined with a singularist theory of causation, contrary to V2, relations between property-types do not ground the causal relation, and hence one cannot be accused of type epiphenomenalism.¹¹ But given a generalist theory of causation, relations between property-types do ground singular causal relations. And it is precisely because anomalous monism assumes such a theory of causation, that is, a nomological theory of causation, that critics consider that it is consistent with the accusation of type

¹¹ This point will be returned to in Part One, § 4.6.

epiphenomenalism. Furthermore, as anomalous monism maintains that only the property types that appear in strict laws are causation grounding and there are no strict psychophysical laws, according to these critics it should be accused of type epiphenomenalism.

2.3 Why Davidson is not a Type Epiphenomenalist

Various attempts have been made to save anomalous monism from type epiphenomenalism. One suggestion is that NCC needs amending — causation need not be backed by strict laws. Hence, LePore and Loewer (1987), attempt to reinstate the mental by appealing to non-strict laws in terms of counterfactual-supporting generalisations hedged by *ceteris paribus* clauses. Alternatively, Smith (1984a, 1984b) argues that V2-epiphenomenalism ceases to be an undesirable position if it is recognised that the way that the mental makes a difference just is that mental events are identical with physical events whose physical properties have a causal role. But these solutions miss the point, for they fail to recognise that the question of whether mental properties of events ontologically ground causal relations, is simply an illegitimate one for Davidson.

In applying V2 to anomalous monism, it is implicitly assumed by both the critics and defendants of anomalous monism, that anomalous monism is committed to a certain understanding of the nature of events as the causal relata. To raise the question of whether events are causes in virtue of their mental properties, or, more specifically, whether mental types ontologically ground causal relations between token events, one must assume a 'Property Theory of Events'. In such theories the ontological category of properties is assimilated to the category of events. Kimean events provide one such example; according to Kim an event is a substance exemplifying a universal at a time. It follows that Kimean events support certain descriptions as opposed to others in virtue of the universal that they exemplify. An event is mental, for example, in virtue of it being an exemplifying of a mental universal.

The second assumption, dependent upon, but not entailed by the first, is that events are causally efficacious 'in virtue of' the ontological types that they exemplify. Given a nomological theory of causation, to advance V2, a distinction must first be assumed between a law statement (which concerns sentences or propositions) and laws that are their worldly correlates. This seems plausible. Indeed, even given a mere regularity theory

of laws (which Davidson adopts¹²), one would assume that one must distinguish between the relata of the law statement and what it is that the objective regularities are regularities between, from which it follows that one must say something about what it is in the world that laws relate. Given the generality of laws, the most obvious candidate are property-types. Hence, relations between property types ground the causal relations between particular events.

Finally, given V2, and the premise that an event exemplifies more than one property type, an event need not be a cause in virtue of all of the property types that it exemplifies. While, for example, the direction of the shot makes a difference to the death that it causes, the loudness of the shot does not. This is because, assuming a nomological theory of causation, given the relevant background circumstances, there is no law linking the type 'loudness' with the type 'death'. The property of loudness is, in this case, V2-epiphenomenal.

It is only given these three assumptions that it is legitimate to advance the complaint of type epiphenomenalism, but Davidson's theory of the causal relata leads to the abandonment of each of them. In 'Thinking Causes' Davidson (1993) suggests that his critics fail to take into account the fact that he considers causal relations to be extensional binary relations between events. (p.6). Suggestions that events are causes as they instantiate laws, that mental events cause something qua physical, or that events are causes in virtue of being physical, turns causation into a multi-termed, non-extensional relation. (p. 13). It is true that given Davidson's account, mental events do not cause physical events qua mental, but it is equally true that mental events do not cause physical events qua physical, for given this extensionalist view of the causal relation it makes no literal sense to speak of one event causing anything qua anything. And clearly as Davidson rejects the 'in virtue of' principle, he cannot be accused of type epiphenomenalism.

The standard response is that V2-epiphenomenalism is compatible with causation being an extensional binary relation between events. Certainly, V2 does not require one to deny that causation is a two-termed relation. It only requires one to recognise that properties are essential in grounding these two-termed causal relations.¹³ As Kim argues, to

¹² See Davidson (1980d) where he advances a weak understanding of Hume's view that a causal law covers every causal claim.

¹³ For example see B. P. McLaughlin (1993, p. 33) and J. Kim, (1993h, p. 22) for this argument.

generate a property epiphenomenalism one only needs to 'acknowledge that the causal relation obtains between a pair of events *because they are events of certain kinds, or have certain properties*. How could anyone refuse to acknowledge this — unless, that is, he believed that causal relations were brute facts about events, having nothing to do with the kind of events that they are?' (1993h, p. 22) But in asking this question Kim, along with most of Davidson's other critics, has clearly failed to properly take on board Davidson's theory of events.

Unlike his critics, Davidson does not consider events to have properties (whether these are universals or tropes), because for him properties are not objective aspects of things in the world. And as Davidson rejects a Property Theory of Events, it makes no sense to suggest that events are causes in virtue of their properties and thus accuse him of a property epiphenomenalism. What makes an event mental (or physical) is whether or not it has a mental (or physical) description. And there is no ontological fact about an event that makes its description as mental (or physical) true or false. Hence, it is with good reason that Davidson refers to his position within the philosophy of mind as a monism rather than a physicalism, because for Davidson, events form an ontologically neutral class of entities. Obviously, therefore, Davidson's distinction between token and type physicalism does not reflect that of his critics, for in denying a type reductionism he is not denying an ontological reduction of the mental to the physical, but a conceptual one. If one had to fit Davidson's position within the mental causation debate into those outlined in the introduction it would have to be that of a property eliminativist. However, unlike most eliminativists, Davidson is not rejecting the existence of mental properties because they are mental, but because he rejects the ontological category of properties. It is for this reason that accusations of type epiphenomenalism can be dismissed.

Davidson's use of terms has not helped to make this fundamental difference between his ontology and that of his critics transparent, because he repeatedly refers to the term 'property' in his discussion of events.¹⁴ However, his use of the term 'property' is not to be interpreted in an ontological sense. For Davidson, talk about properties is simply talk about the predicates that can be ascribed to an event when the event is variously described. It does not entail an ontological commitment to properties. Davidson's use of the term 'law' invites similar confusion. Although Davidson is committed to an ontological theory of causation, he does not make a distinction between law statements and laws. By the term 'law' he means a law statement, and thus 'laws' connect events by connecting

¹⁴ For example, see D. Davidson (1993, p. 7 & p.13).

predicates, not by connecting properties. Hence, Davidson's claim that 'events instantiate a law only as described in one way rather than another...' (1993, p. 6).

The critics who accuse Davidson of type (or indeed trope) epiphenomenalism have failed to understand his theory of the causal relata, basing anomalous monism within a property theory of events. Kim argues that 'The issue has always been *the causal efficacy of properties of events...*' (1993h, p.21). This may be the case given Kim's own theory of events, but one cannot embed anomalous monism in whatever theory of events one likes. Certainly, V2 would generate a plausible criticism of anomalous monism if it was embedded within a Kimean theory of events, but to criticise Davidson's theory under a scheme of event that is not his own would be question-begging. Anomalous monism and Davidson's theory of events must go hand in hand. And, quite clearly, one cannot accept Davidson's theory of events and then go on to accuse Davidson's anomalous monism of type epiphenomenalism.

Surprisingly, this has not always been clear to Davidson's critics. Of those philosophers who charge Davidson with type epiphenomenalism, some do mention his nominalism but treat it as unimportant to the discussion. For example, Baker comments that Davidson and Kim have a different understanding of the term 'event', but states that she will assume a 'hybrid' use of the term when discussing Davidson. (Baker, 1993, p. 75, fn.1) Her assumption is that in order to talk about property epiphenomenalism it does not really matter whether one understands 'property' in an ontological sense or a linguistic one. McLaughlin (1989) makes this assumption more explicit. According to him, the criticism of Davidson can be recast without appealing to ontological properties. As he puts it: 'The Principle of Causality implies that events can participate in causal relations only in virtue of satisfying strictly nomic descriptions. Anomalism implies that no mental description is strictly nomic. So the Principle of Causality and Anomalism imply that no event can participate in a causal relation in virtue of satisfying a mental description.' (1989, p. 122). However, I fail to see how one can do any such thing. Given Davidson's understanding of the term 'property', V2 is clearly false. There is no question of predicates making a causal difference. Predicates do not ground a causal relation. Whether an event is described in one way rather than another does not make a difference to the causal effects that it has. As Davidson (1993) states: 'If causality is a relation between events, it holds between them no matter how they are described' (p.6). 'Naming the American Invasion of Panama 'Operation Just Cause' does not alter the consequences of the event' (p.8). If one event

causes another, it causes it regardless of how it is described. One will only be tempted to deny this if one assumes that predicates have ontological correlates.

2.4 Predicate Epiphenomenalism

Let us consider one final interpretation of the 'in virtue of' principle. As noted earlier, in 'Nonreductive Materialism and Mental Causation', Marras argues that what is really at issue 'is not so much the causal *efficacy* of mental properties as rather the *relevance* of mental properties to the formulation of causal explanations.' (1994, p. 473) Marras distinguishes between event epiphenomenalism, which is concerned with the causal efficacy of mental events, and property epiphenomenalism, which is concerned with the causal relevance of mental properties. His concern is whether, given anomalous monism, mental properties can have any causal relevance within causal explanations of physical events, or whether it is the case that, for example, although my desire for water did cause me to get a drink, it's being a desire for water was causally irrelevant to explaining my drinking (p. 474).

It is indeed questionable whether Marras is intending to offer a form of epiphenomenalism distinct from type epiphenomenalism. Certainly, if the two positions are distinct, then, given a nomological theory of causation and the covering law theory of explanation, they are intimately related. According to the nomological theory of causation, for two events to be causally related is for two property types of the events to be nomologically related. And given Hempel's deductive nomological model of explanation, deterministic explanations, including causal explanations, are deductive nomological arguments, where deductive nomological arguments are sound arguments that include essentially as a premise some law. Thus to explain an event by citing its cause is to identify an event pair as an instance of a lawful relation between event types. Thus Marras may simply be drawing our attention to the fact that causal explanation demands not only that an event is singled out as a cause, but that furthermore, the salient property of the cause is identified.

However, the claim that mental properties are causally redundant within psychophysical causation and the claim that mental properties are causally redundant within psychophysical explanation should not be conflated. Causal relata are ontological entities, therefore the first claim requires an ontological understanding of the term 'property.' Causal explanantia are linguistic entities, therefore the latter claim requires a linguistic

understanding of the term 'property'. Unfortunately, Marras repeatedly conflates the ontological and linguistic understandings of the term.¹⁵ Consequently, he fails to adequately distinguish between the claim that properties (in the ontological sense) are the things that ground the causal relationship in virtue of the objective nomological relations that hold between them, and the claim that properties (in the linguistic sense) are the linguistic entities that causal explanations relate and which predicates are of causal relevance within our descriptions depends upon which predicates our law statements relate. This leads to the general conflation of causation with causal explanation.

In so far as Marras' concern is with the first claim, he has fallen into the same error as McLaughlin, Kim et al. — that of failing to take into account Davidson's nominalism. But what of the alternative claim, which I shall refer to as predicate epiphenomenalism? Its concern is with the causal relevance of mental predicates within psychophysical explanations. It is uncontroversial that sentences containing mental descriptions successfully explain physical events, rather, the point is that these explanations are redundant. The criticism is that anomalous monism leads to the conclusion that it is a mental event falling under a physical description, not a mental description, that causally explains events within the physical domain. This is because, for Davidson, who adopts the deductive nomological model of the explanation, the explanatory force of a claim is derived from the underlying causal law. According to AM, there are no strict psychophysical laws linking mental concepts with physical concepts. It is therefore not mental events qua mental that causally explain physical events.

As it does not assume the existence of properties, predicate epiphenomenalism is compatible with Davidson's nominalistic ontology, and therefore to this extent it is legitimate to accuse anomalous monism of it. Furthermore, according to Davidson, while causation is an extensional relation, causal explanation is an intensional one.¹⁶ For Davidson, laws are linguistic entities and thus whether an event instantiates a law, and thus can be explained by a law, depends upon the way it is described.

But is it thus plausible to accuse Davidson of predicate epiphenomenalism? Surely not, for anomalous monism does not require that in order for mental explanations to be causally

¹⁵ His assumption that, for example, given type physicalism mental properties would be causally efficacious would suggest that he understands the term in an ontological sense (1994, p. 470). His appeal to properties as the causal explanantia would suggest that he is interpreting the term in a linguistic sense.

¹⁶ See D. Davidson (1980c).

relevant in physical explanations, there must be a strict psychophysical law statement relating them. It is compatible with AM that there are non-strict psychophysical law statements. Indeed, Davidson explicitly claims that there are such law statements. (See, for instance Davidson (1993, p. 9)). Furthermore, anomalous monism only requires that all causal relations are backed by strict law statements, not that all causal explanations include strict law statements.

Given anomalous monism, it is indeed true that where there is a causal explanation of a physical event in mental terms there will also be a causal explanation of the event in physical terms, but this does not make the mental explanation in some way causally irrelevant. (See Davidson (1993, p. 16)). The thought that it does may arise if one has Kim's principle of explanatory exclusion at the back of one's mind when considering anomalous monism. (This states that one cannot have two complete and independent causal explanations for the same event). But Kim's principle of explanatory exclusion gains its motivation from accepting explanatory realism, roughly the principle that a causal explanation of *e* in terms of *c* is a correct explanation only if *c* is in reality a cause of *e*. While the principle of explanatory realism, and thus the principle of explanatory exclusion, is perhaps plausible given Kim's ontology, it is certainly not given Davidson's.¹⁷ Explanation is a relation between statements, and there is, according to Davidson, no thing that makes a statement true.¹⁸ Given that there is nothing in events that explains why events support certain descriptions as opposed to others, there is no reason to suggest that one could not have two complete and independent explanations of the same event. Anomalous monism is compatible with the possibility of predicate epiphenomenalism, but it certainly does not motivate such a position.

2.5 The Rejection of Davidson's Theory of the Causal Relata

Given a Davidsonian theory of events, accusations of epiphenomenalism are clearly inappropriate. Davidson's rejection of the ontological category of properties means that his position is incompatible with V1 and V2, and hence, he cannot be accused of either V1 or V2 epiphenomenalism. Although Davidson's ontology is consistent with predicate epiphenomenalism, it certainly does not provide a threat to anomalous monism. Hence

¹⁷ For Kim's Principle of Explanatory Exclusion and his defence of explanatory realism see Kim (1990) and Kim (1993e). In Part 2, § 4.3. I argue that this principle is implausible even if one accepts Kim's ontology.

¹⁸ To quote Davidson: 'Nothing... no *thing*, makes sentences and theories true...' (1984b, p. 194)

Davidson cannot be accused of property epiphenomenalism, regardless of whether one interprets 'property' in an ontological or a linguistic sense. However, there is a fundamental difference between the way in which Davidson and his critics within the philosophy of mind approach causation and the causal relata. To see what is wrong with anomalous monism, one must go back a step and consider Davidson's theory of the causal relata. It is the implausibility of Davidson's theory of the causal relata that leads to the rejection of anomalous monism.

It is one of the most commonly assumed claims within the philosophy of causation that causes have their effects in virtue of certain of their properties. Given a nomological theory of causation, it is natural to assume a distinction between a law and a law statement, the first being a relation between properties, the second a relation between predicates. Causation is a relation between laws, hence relations between property types ground the causal relations between events. But as we have seen, given Davidson's theory of the causal relata, this is not the case. Causes, according to Davidson, do not have their effects in virtue of properties. For Davidson, there are only law statements relating predicates. There is no ontological correlate, a relation between properties. Thus NCC amounts to the claim that all singular causal relations are backed by strict statements of law. Given AM there are no strict psychophysical law statements. And on this basis, given CI, anomalous monism concludes that mental events are identical with physical events.

Davidson's nominalism stems from the fact that his primary interest is in a semantics for natural language. His ontological commitments are determined by his semantical stance, and according to him semantics can do just as well without properties. In considering the semantical structure of event statements, Davidson concludes that one must admit the ontological category of events (and actions) as values of bound variables.¹⁹ But following Quine, he maintains that the predicate of a true statement in canonical form ought to harbour no ontological commitments, ontology only being carried by the term that is subject to quantification. Hence, for example, the predicate 'is red' is true of the event to which it applies, but the question of what makes such a predicate apply to this event and not others should be rejected.

Davidson's semantical stance determines not only which ontological categories he believes to exist but also his discussion of causation. Distinguishing the question of the logical form of causal statements from the question of how the causal relation should be

¹⁹ For a defence of this point see Davidson (1967a, 1967b and 1980c).

analysed, Davidson's concern is with the former.²⁰ Analysis of the causal relation is the analysis of a two-place predicate in an ordinary extensional first-order language, and analysis of the logical form of causal sentences shows that causation is a relation between events. But Davidson does not attempt to provide an analysis of the causal relation itself. Indeed, consideration of the logical form of causal statements does not give us any reason to quantify over causal relations, only causes and effects. But despite these considerations, one of Davidson's central premises within anomalous monism is that causal relations must be backed by strict law statements, and in assuming this, Davidson is clearly taking a stance on the analysis of the causal relation. However, Davidson never advances any arguments for NCC and it is a position that rests uneasily with his nominalism.

Davidson claims that all causal relations must be backed by strict law statements. But how is one to analyse a semantic type given nominalism? If there are no properties, then one faces the problem of what it is that marks off sets of events whose members are more or less alike from sets whose members are not alike.²¹ This raises the question of how Davidson is to construct classes that will yield genuine regularities, for without properties it becomes very hard to give any content to the notion of an objective regularity.

More generally, if all causal relations are backed by strict law statements then surely there must be a fact of the matter about which law statements are true, that is surely there must be something in the world that makes law statements true. But if so, then surely this must be an ontological relation between property-types. Hence, to maintain a nomological theory of causation one must admit the ontological category of properties.

This last line of reasoning assumes some form of truthmaker principle, according to which for every contingent truth there must be something which ontologically grounds it.²² Indeed, acceptance of the truthmaker principle leads to the rejection of the view that predicates do not have to be taken seriously in considering the ontological implications of any true statement, not just true law statements. The statements 'The apple is red' and 'The apple is round' are true of the same particular. Given the truthmaker principle there must be some ontological ground which accounts for the difference between applying the predicate 'red' to the apple and applying the predicate 'round' to the apple. It is something

²⁰ See Davidson (1980c).

²¹ See Armstrong (1997, p. 221) for this criticism of nominalism.

²² For a defence of the truthmaker principle see Armstrong (1997).

ontologically different about the particular that makes each statement true. This ontological difference is accounted for by appealing to the ontological category of properties.

But although Davidson obviously considers law statements to have a truth-value, he does not raise the question of what it is that makes these law statements true. Nor need he do so, because he is not committed to the truthmaker principle. The truthmaker principle is attractive if one maintains a correspondence theory of truth because the latter tries to explain what it is for a sentence to be true in terms of its correspondence to something else. However, Davidson's advocacy of events as particulars and his rejection of events as facts (Davidson employs the Fregean argument that there is only one fact) underlies his objection to the correspondence theory of truth of the type advanced by Russell (1918) and Wittgenstein's *Tractatus*.²³ Davidson advances a Tarskian semantic theory of truth, which, unlike the correspondence theory of truth, does not provide an explanation for the truth of a statement. For Davidson, the things that correspond to statements are not states of affairs but the events that satisfy them, but the correspondence between a statement and an event does not explain the truth of that statement. Hence, Davidson does not need to appeal to properties to explain why events support certain descriptions rather than others because, according to him, there need not be anything about events that explains why events support certain descriptions rather than others.²⁴

Some would find Davidson's claim that there is nothing that makes a statement true objectionable, maintaining some form of truthmaker principle. Given such a position there must be some relation in the world that makes causal statements true. Given a nomological theory of causation, it is the fact that there is a nomological relation between types, combined with the fact that these types are instantiated by particulars. But even given doubts about the truthmaker principle, it is most plausible that Davidson has got things the wrong way around. One's motivation for accepting or rejecting an ontological category and hence a theory of the causal relata should not have semantic considerations at its base, because contrary to Davidson, a theory of meaning cannot be appealed to to settle ontological issues. An appropriate semantics is to be construed only after the ontological issues are settled. Hence, whether or not one should admit the ontological category of properties is not to be based upon consideration of whether sentences require quantification over properties, and whether those that do can be paraphrased to avoid any

²³ For Davidson's version of the 'slingshot' argument see Davidson (1984c).

²⁴ For example, see Davidson (1984b, p. 194).

such commitment. Rather, whether one should admit the ontological category of properties, and indeed what properties are, is to be established through metaphysical enquiry. That is, by consideration of whether such a category could exist (and thus their existence and identity conditions), whether they play a non-redundant role in one's ontological system, and whether they can co-exist with the other ontological categories that one includes within this ontological system. As causation is an ontological relation, what it is a relation between is to be established through consideration of these ontological categories along with consideration of the relation of causation itself. If one approaches ontology in this way, then properties will inevitably play an essential role within one's ontological system and more specifically within one's theory of causation. It is for this reason that Davidson's theory of the causal relata should be rejected and, along with it, anomalous monism.

3

Coarse-grained Theories of the Causal Relata

3.1 Coarse-grained Theories and the Qualitative Specificity of Causation

The ontological category of properties should be admitted; there are entities that characterise objects. Objects are mental if characterised by mental properties and physical if characterised by physical properties. Furthermore, regardless of what the causal relata are, they must be related to the ontological category of properties, for if there are properties, then they play the central role within causation: the causal relations that an entity can enter into depend upon what that entity is like, and hence which properties characterise it. Exactly how this is to be spelled out will depend upon the theory of causation under consideration. But regardless of whether what matters to causation is the entities with which another is constantly conjoined or nomologically connected, or the powers it possesses, or the energy it transfers, what is crucial in each case is the properties that characterise that entity. Whether a theory of causation requires that properties play any further role in addition to being characterising entities depends upon the theory of causation under consideration. Hence generalist theories require that token causal relata exemplify types such that at the type-level the entities are suitably connected. This follows if, as well as being characterising entities, properties (or more questionably sets of properties) play a unifying role.

But although properties play a central role within causation, the theory of the causal relata that Davidson's critics mistakenly assume Davidson to adopt is unsatisfactory. In this section, I shall argue that the combination of the qualitative specificity of causation and the transitivity of causation, means that causes cannot have epiphenomenal properties. This means that for mental causes to have physical effects, mental properties must make a

causal difference within the physical domain, which in turn allows us to provide a straightforward articulation of the argument from causal overdetermination.

So far, our discussion has focused upon coarse-grained theories of the causal relata. Critics of anomalous monism assumed Davidsonian events to be qualitatively dense, that is, for an event to instantiate numerous properties and for it not to be identifiable with an instantiation of any one of these properties. Contrary to this, Davidsonian events in fact lack any grain at all. But if one were to maintain that, for example, an event is constituted (in part) by all the properties that are instantiated within a particular spatio-temporal zone this would be to adopt a coarse-grained theory of the causal relata. The theory of events proposed by Cynthia Macdonald (1989) in her discussion of mental causation is also coarse-grained.²⁵ According to it, events are the causal relata, where events are property-instances. Macdonald's property-instances are constituted by the categories of substance, property and time, and hence in this respect are like Kimean events. However, unlike a Kimean event, according to Macdonald a single property-instance may have more than one property and consequently is coarse-grained. Alternatively, to maintain that the causal relata are substances (where substances are things like tables and stones), is to maintain a coarse-grained theory of the causal relata.

Consideration of the causal relation shows that the causal relata must possess a degree of qualitative specificity. In every causal relation it is specific property-types that ground the causal relation, or for singularist theories of causation, specific property-instances that make the causal difference. Due to the qualitative specificity of causation, it is highly improbable that all of the properties of a coarse-grained entity will play a causal role — as well as causal properties, the causal relata will have epiphenomenal ones. Coarse-grained theories must therefore supplement their theory of the causal relata with an 'in-virtue of principle in the form of V1 or V2. Where *c* and *e* are a cause and an effect, it is a specific property of *c* and a specific property of *e* in virtue of which *c* and *e* are subsumable under a law, or in virtue of which if *c* had not occurred *e* would not have occurred, or in virtue of which *c* raises the chances that *e* will occur, etc. So given the 'in virtue of' principle, although the causal relata are entities characterised by multiple properties, one can deny that every property that characterises it, makes a difference to the causal relation that it enters into.

²⁵ Also see Cynthia and Graham Macdonald (1986) and (1991).

As we have seen, given this kind of theory of the causal relata, even if mental causes have physical effects (in virtue of being identical with physical causes), it is legitimate to ask whether the mental properties of the cause are epiphenomenal, for mental causes may not be causes in virtue of their mental properties. Thus, for example, if substances are the causal relata, and mental substances have physical effects because both mental and physical properties are possessed by the body, if the mental properties of the body make no causal difference to the physical properties of the body, then the mental properties of a mental cause will be epiphenomenal. For this reason, I do not think that Macdonald's claim that 'an instancing of a mental property by an event just is an instancing by it of a physical one' (1989, p. 162) offers a response to the problem of mental causation. Given her theory of the causal relata, mental and physical property-instances can be identified, despite the fact that mental and physical properties are distinct. But this does not resolve the problem of mental causation, for even if mental and physical properties can be instantiated in a single property-instance, the qualitative specificity of causation requires one to consider whether it was the mental or physical property of the property-instance that made the causal difference, and having raised this question, one must face the same dilemma faced by all other property-dualists.

3.2 The Transitivity of Causation and the Rejection of Coarse-grained Theories

Coarse-grained theories of the causal relata are, I would suggest, unsatisfactory. This is because causation is a transitive relation, but if a coarse-grained theory is supplemented with an in virtue of principle, this leads to violations of transitivity.²⁶

The problem arises when the property of the causal relatum in virtue of which it is a cause, is not the same property in virtue of which it is an effect. Hence, suppose that e1, e2, and e3 are coarse-grained events, that P1, P2, P3 and P4 are properties, and that e1 has P1, e2 has P2 and P3, and e3 has P4. Furthermore, let us assume that there is a causal relation between e1 and e2 in virtue of e1 having P1 and e2 having P2, and that there is a

²⁶ Also see Ehring (1997, ch. 3) who raises a similar argument against coarse-grained theories of the causal relata, but mistakenly targets it against Davidson's theory of the causal relata. For the purpose of this discussion, I shall assume that causation is transitive. However, this feature of causation is not beyond question. See Lowe (1980) for an argument against the transitivity of causation.

causal relation between e2 and e3 in virtue of e2 having P3 and e3 having P4. In order for this to be consistent with the transitivity of causation, e1 must cause e3. However, contrary to this, there need not be any causal relation between e1 and e3, in virtue of e1 having P1 and e3 having P4. For example, let us say that e1 is the event of putting red dye in a glass of water, e2 is the event of there being red water in the glass, e3 is the event of the ice-cubes melting (which have been placed in the water after it is red). Normally one would want to say that e1 causes the redness of the water, but given a coarse-grained view of events, one should in fact say that e1 causes the event of there being red water in the glass. That is e1 causes e2. Equally, given a coarse-grained view of events, e2 causes e3. Given the transitivity of causation, it should therefore follow that e1 causes e3. That is the event of placing red dye in the water causes the ice cubes to melt. But this is false. This can be seen by appealing to the counterfactual consideration that, if the water had not been red, it would still have caused the ice to melt. (Clearly this failure of counterfactual dependence does not arise because it is a case of pre-emption or overdetermination.)

One might respond by denying the conclusion of the argument. The event of placing red dye in the water was the cause of the ice cubes melting in the red water. It is true that the melting of the ice would have occurred if the first event had not occurred, but it would not have been the event of the ice melting in red water. That is, the effect e3 would have lacked a property if it was not causally linked to e1, so e1 does make a difference to e3 in the sense that e3 would have been different if e1 had not occurred. However, there remains an obvious sense in which the water being red plays no causal role in the ice melting, it has no bearing on e's being a melting. Indeed, the response suggests that every property of an event is essential to it, but such an essentialism of events seems implausible.

To avoid violating the transitivity of causation, the causal relata cannot have properties that are both causal and non-causal within a single chain. One is thus left with two responses. The first is to hold on to a coarse-grained theory, but rule that the property in virtue of which a coarse-grained entity is a cause must be the same property in virtue of which it is an effect. That is, in a single causal chain, only one property of a coarse-grained entity can make a causal difference. But this is either false (for causal situations such as the above regularly arise) or it is to put an artificial restriction on the coarse-grained entities that can qualify as the causal relata — the causal relata that one is left

with may well be coarse, but they will be a highly selective group of entities that have been picked out for no further reason than that they do not lead to violations of transitivity.

The more plausible alternative is to abandon any theory of the causal relata that needs to appeal to the in virtue of principle, and hence maintain that a cause can only have causal properties. Such fine-grained views of the causal relata do not encounter problems with transitivity, because they will pick out different entities where a coarse-grained theory will pick out only one. Given, for example, the property-exemplification view of events, according to which an event is the exemplification of a property by a substance at a time, the redness of the water in the glass at t1, and the temperature of the water in the glass at t1 are two different events. Putting red dye in the glass of the water causes the first event, but it does not cause the second and it is the second event that causes the ice to melt. There is no violation of transitivity because there is not a causal chain from the redness of the water to the melting of the ice.

To summarise, due to the qualitative specificity of causation, coarse-grained theories of causation are incorrect. Their mistake was to start with the idea that a thick-grained entity such as an event or substance was the cause and then, due to the qualitative specificity of causation, 'work in' by supplementing their claims with an 'in virtue of' principle. To avoid violating transitivity, the proper method is, I would suggest, to start by identifying the causal property (or properties) in a causal relation and then 'work out' from the relevant property, specifying what the instantiation of this property is an instantiation by.

Of course, this is only to reject a coarse-grained theory of the causal relata; it is not to reject a coarse-grained theory of events. Nor does it mean that coarse-grained events (or substances) are no longer of significance within causation, for the causal relata may be an aspect or cross-section of such an entity.²⁷ Such an aspect has an ontological status and is not the product of a discriminating mind. However, it is 'abstract' in the Lockean sense that it is only possible for it to exist in the world in conjunction with other such aspects, although in thought it may be separated from them. Hence, for example, although in thought one can separate the ball's redness from any other state of the ball, in actual fact the ball must also have a shape, a size, etc. If aspects of events (or substances) were the causal relata, this would indicate that those who consider the causal relata to be coarse-

²⁷ See Honderich (1988) who maintains that the causal relata are aspects of substances and Dretske (1977) who maintains that the causal relata are aspects of events.

grained entities have confused the thing that the cause is a part of (the whole event or substance) with the cause (an aspect of an event or substance).

Finally, note that these considerations only rule out epiphenomenal properties of the cause. They do not entail that there are no epiphenomenal properties. The aspect view, for example, does not rule out epiphenomenal properties of coarse-grained events or substances. This is precisely because it will only be an aspect of the event or the substance that is the cause.

3.3 A Re-articulation of the Argument from Causal Overdetermination

These considerations allow us to provide a straightforward articulation of the argument from causal overdetermination. Causal specificity requires a fine-grained interpretation of what it is to be a cause, and hence, what it is to be a mental cause, because to appeal to some sort of 'in virtue of' principle leads to violations of transitivity.

As causes do not have epiphenomenal properties, if a cause is a mental cause, that is, if a cause is characterised by a mental property, and the mental cause is required to play a causal role in the physical domain, then the mental property must make a causal difference within the physical domain. Unless mental causes just are physical causes (in which case, mental properties must be physical properties), given the denial of systematic causal overdetermination, this entails the rejection of the causal closure principle.

The argument from causal overdetermination is not an argument that can be raised at increasingly finer levels. The question within the mental causation debate is simply whether there can be mental causes (within the physical domain), that are not identical with physical causes. If there are mental causes within the physical domain, there can be no further question of whether 'the mentalness' of the mental cause is in some sense causally redundant within the physical domain.²⁸ Therefore, contrary to Yablo (1992), the argument from causal overdetermination does not raise two problems for mental causation, 'one about mental particulars (events), the other about mental properties.'

²⁸ Although this is to assume that properties do not themselves have aspects — a point to which I shall return later in this discussion.

(1992, p. 248) The argument from causal overdetermination raises only one problem for mental causation. It is one about mental causes.

4

The Properties of Causation and the Mental Causation Debate

4.1 Property-instantiations

Section Two concluded that the causal relata have properties. Section Three concluded that the causal relata do not have non-causal properties. What then are the causal relata? Making the ontological distinction between a substance and a property, where substances are particulars that instantiate properties but cannot themselves be instantiated, the causal relatum is a substance's instantiation of a property, or what I shall refer to as a property-instantiation.²⁹

If the causal relata are property-instantiations, then the argument from causal overdetermination is as follows:

1. A substance's instantiation of a mental property has a physical effect
2. Every physical effect has as a sufficient cause a substance instantiating a physical property
3. There is no systematic causal overdetermination.

The conclusion of the argument from causal overdetermination is that a substance's instantiation of a mental property is identical with a substance's instantiation of a physical property. In order for two property-instantiations to be identical, they must involve the same substance and property, hence in order for mental causes to be identical with

²⁹ Here it may be objected that one should identify the causal relata not with the having of a property by a substance, but the changes in the properties of a substance. But although causation and change are intimately related (a cause implies a change in an effect, and in order for the effect to be brought about there must have been a change in either the intrinsic or relational properties of the cause), it does not follow that the causal relata must themselves be changes. Rather, it is up to a theory of causation to capture this notion of change.

physical causes, mental substances must be identical with physical substances and mental properties must be identical with physical properties.³⁰

Now at least for the purpose of discussions of the causal relata, to simply say that the causal relata are property-instantiations is not enough. To be interesting, a theory of the causal relata must consider what the ontological categories of property and substance to which property-instantiations are related actually are. In the first place, what are the properties of causation? That is, if the causal relata are property-instantiations, are the properties that substances instantiate universals (e.g. Armstrong (1989) and (1997), Fales (1990)), or particulars (e.g. Campbell (1991), Williams (1966), Martin (1980)) And what is the ontological status of substances? Is this ontological category fundamental, or are substances nothing over and above bundles of properties? And there are further related disagreements about what a property-instantiation actually is. For example, Armstrong (1997) recognises both substance and property to be fundamental, but considers them to be constituents of states of affairs, hence identifying property-instantiations with states of affairs. Alternatively, although maintaining a two-category ontology one could identify a property-instantiation not with a state of affairs, but with a state of a substance. These different ontological systems generate very different understandings of a property-instantiation, not all of which are equally plausible candidates for the causal relata.

But does any of this matter to the mental causation debate or can a neutral stance be taken amongst these various understandings of a property-instantiation? That it does not matter would certainly appear to be the general assumption within the mental causation debate. Although property-instantiations are often appealed to as the causal relata, the question of how to understand the ontological categories to which property-instantiations are related tends not to be considered. Nor is it generally thought to matter whether property-instantiations are states of affairs or states of substances. Indeed, discussions often switch between talking about facts as the causal relata and talking about events as the causal relata, the underling assumption seemingly being that at least for the purpose

³⁰ Note, therefore, that if properties are universals, this understanding of the causal relata does not allow for an interesting distinction between token and type physicalism. If a mental universal is involved in one property-instantiation and a physical universal is involved in another, and the mental and physical universals are distinct (hence type dualism), then the mental and physical cause must be distinct (hence token dualism).

of the mental causation debate, it does not matter whether the causal relata are events or facts.³¹

Contrary to this, I consider that these issues matter immensely to the mental causation debate. I shall argue that the plausibility of the argument from causal overdetermination and the possible responses to it, greatly depend upon what the properties of causation are — that is, whether they are universals or particulars. But what the properties of causation are, depends upon whether the causal relata are states of affairs or states of substances, which in turn depends upon what substances are. Consequently, as those within the mental causation debate need to take a stance on what the properties of causation are, they need to take a stance on whether the causal relata are states of affairs or states of substances, and hence what substances are. In § 4.3 - § 4.6 I consider how one's understanding of the nature of property — that is, whether properties are particulars or universals — affects the mental causation debate. In § 5, I argue that the properties of causation must be particulars rather than universals. But first, let us consider what these two competing theories of properties are and how they differ.

4.2 Property-instances

If the causal relata are property-instantiations, then what are the properties that substances instantiate? One understanding of a property is as a universal. The proponent of universals maintains that properties can be multiply exemplified by wholly distinct substances at the same time. Thus, for example, if the red of this apple and the red of my dress are both the same shade they exemplify the same universal, where sameness is to be understood in a strict sense; as a genuine identity. There is, in other words, a single property 'redness' which the dress and the apple literally share. If properties are universals, then as well as characterising substances, properties are unifying entities; in answer to the question of what it is about a substance in virtue of which it is red at time *t*, and what it is about these two substances in virtue of which they are both red at *t*, the proponent of universals can give the same answer — it is in virtue of their having the universal 'redness.'

³¹ For example, see LePore and Loewer (1987) and Heil (1992, p.16) who both conflate states of affairs with states of substances.

Singular causation is a relation between entities existing at the level of particulars. Hence, the causal relata cannot be universals. But while universals do not exist at the level of particulars, exemplifications of universals do. Although, for example, the universal property of redness is exemplified at the same time by red substances s1, s2 and s3 and is strictly identical within these exemplifications, the exemplifications of red in s1, s2 and s3, that is the red of s1, the red of s2 and the red of s3 are distinct particulars, the particularity deriving not from the property, but the particularity of the substance which the property is exemplified by. Hence, what is particular about the red of this apple, is that it is the red of this apple as opposed to the red of anything else. There is nothing about the redness itself that makes it a particular. Fitting this into the property-instantiation account of the causal relata, it is, for example, the apple exemplifying the universal property of redness at time t that causes me to pick it, where the property-instance is here to be understood as this particular exemplification of the universal redness.

Exemplifications of universals are not the only kind of property-instance. Instead, one could understand properties as characterising particulars. I purposefully do not use the term 'trope' because tropes are typically associated with a bundle theory. By the term 'characterising particular', I understand something that is neutral between a one-category and a two-category ontology. (Hereafter, I shall refer to a characterising particular as a 'CP'). Unlike a universal, a CP is a particular and hence cannot be instantiated by distinct substances at the same time. Even if they exactly resemble each other in respect of their colours, my dress and this apple do not literally share the same CP; the dress's redness and the apple's redness are numerically distinct.

Unlike universals, as they are not multiply instantiable, CPs are not unifying entities. But despite this fact one can explain the instantiation of types without recourse to universals, most plausibly in terms of resemblance.³² At the type level, one has sets of resembling CPs, two CPs being of the same type if and only if they belong to the same resemblance class. Hence one can say, 'The red of my dress and the red of this wine are the same', but here sameness does not imply a strict identity, but instead a relation of resemblance. Each red thing has its own property of redness, but all red things resemble each other more or less closely.

³² D. C. Williams (1966) and Campbell (1991) combine a trope theory with an account of the unity of tropes in terms of resemblance.

To gain a better understanding of the difference between a CP and a universal, one requires a clearer understanding of what it is for a property to be a particular. One proposal is to appeal to spatiotemporal considerations, defining a CP as a property that can only be wholly present in one spatial location at a time, and a universal as a property that can be wholly present at different spatial locations at a time. Certainly, this is often appealed to, to introduce the notion of a particular property.³³ However, one should resist defining a particular property in spatio-temporal terms. It would be plausible, given the assumptions that all properties are located wherever the substances that instantiate them are and that all substances are in space-time. For it would then follow that CPs would have a single location in space at a time, namely the spatial location of the individual substance that instantiates them, and universals being instantiated by more than one substance would exist in more than one spatial location. But as we have seen, one cannot rule out by definition the possibility that some substances lack a spatial location. And it is only if all properties are necessarily instantiated by spatial substances, that one could define a particular property as that which possesses a single location in space at a time.

A better suggestion is simply that a particular property is a property that cannot be instantiated by more than one substance at a time and a universal is a property that can be.³⁴ This does require one to adopt the Aristotelian position that properties cannot exist uninstantiated, but given a proper understanding of the nature of properties, this is plausible. And as it is not being assumed that for a property to be instantiated it must be instantiated by a spatial substance, one avoids begging the question against those ontologies that propose non-concrete substances. Equally, if one was to assume an in re conception of properties, this understanding of a particular property would explain why CPs cannot have more than one spatial location at a time while universals can; a substance does not possess more than one location in space at a time, and therefore as a CP cannot be instantiated by more than one substance at a time, it too can only possess one spatial location at a time. Universals are instantiated by more than one substance at a time, and hence given immanent realism, will be wholly present in more than one spatial location at a time. Hence, in the following I shall understand a CP to be a property that can

³³ For example, see Loux (1976, p.4) and Robb (1997, p. 186)

³⁴ Armstrong (1997, p. 21) seems to adopt this understanding of a particular property. Note, that here I am understanding the term 'substance' in a metaphysically neutral sense.

only be instantiated by one substance at a time, and a universal to be a property that can be instantiated by more than one substance at a time.³⁵

Before moving on, a terminological point needs to be made about the relationship between a substance and a property. For want of a better term, I shall refer to the relationship between a substance and a property as one of 'instantiation.' (Hence my claim that the causal relata are property-instantiations.) By 'instantiation' I intend a relationship, which is neutral between the understanding of a property as a CP and a universal, and thus neutral about whether the property-instance is the property itself or the instantiation of the property. Thus if the apple instantiates the property of redness, the property-instance may be either the instantiation of the property of redness or the property of redness, depending upon whether the property of redness is a universal or a CP.

4.3 Assumptions within the Mental Causation Debate

When philosophers refer to property-instances, and no prior stance has been taken on the universal-particular debate, it is often difficult to decipher whether they are understanding property-instances in terms of instantiations of universals or CPs.³⁶ Indeed, some think that as far as the discussion of the causal relata is concerned, one can maintain a neutral stance. For example, in his discussion of the causal relata Hausman (1998, p. 18) refers to property-instances as 'tropes' but purposefully does not distinguish tropes from exemplifications of universals. I would suggest that such a neutral stance cannot be taken. In § 5, I will argue that it is only given an understanding of property-instances as CPs, that one can advance a plausible theory of the causal relata.

Crucially for this discussion, neither can the mental causation debate be neutral between the two understandings of a property-instance, for what property-instances are greatly affects the plausibility of certain responses to the argument from causal

³⁵ It might be objected that this criterion of particularity fails because it is only possible for some universals to be instantiated once. For example, even if 'being the largest piece of gold on earth' is a universal, 'being the largest piece of gold on earth at time t' can only be instantiated once. But this is a peculiar feature of one type of dyadic universal. Certainly, all intrinsic universals can be instantiated more than once.

³⁶ For example, see Honderich's (1988, p. 15) discussion of the causal relata and property-instances.

overdetermination. Before we consider the reason why property-instances must be CPs, let us consider why this conclusion will be so important to the mental causation debate.

Although rarely made explicit, the underlying assumption within the mental causation debate is that properties, and hence the entities that make a causal difference in the physical domain, are universals. That those within the mental causation debate commonly assume that the properties of causation are universals, can be seen by the general assumptions that are made about properties within the mental causation debate. For example, it is commonly assumed that a property can be shared by different particulars. Hence, consider the way that psychophysical supervenience is standardly formulated. The formulations typically involve the assumption that properties can be had by more than one particular at the same time.³⁷ This clearly can only be the case if properties are universals. (This probably stems from the fact that discussions of supervenience take their lead from Kim, who explicitly assumes that properties are universals). To give another example, the identification of mental and physical properties is assumed to be a type-identification. For example, Marras (1994, p. 465) and Seager (1991, p. 193) both assume type physicalism to be the thesis that mental properties are identical with physical properties. Once again, this is to assume that properties are universals.

As it is being assumed that properties are universals, questions about whether mental properties exist, whether they are independent from, identical with, or supervenient upon physical properties, are questions raised at the level of types. The various positions within the mental causation debate are compatible with a CP ontology. Mentalism, for example, remains largely unaltered if it is articulated within a CP ontology. According to it there are mental CPs and these are neither identical with, nor ontologically dependent upon physical CPs. A substance mentalist and a property mentalist differ on the question of whether mental CPs are instantiated by the body. Given interactive mentalism, mental CPs make a causal difference in the physical domain. On the other hand, the eliminativist denies that there are any mental CPs, whilst the psychophysical reductionist maintains that mental CPs are identical with physical CPs. But just because these various positions within the mental causation debate can be fitted into a CP ontology, one would be mistaken if one was to assume that a shift from universals to CPs would be of no great significance to the mental causation debate.

³⁷ See, for example, Heil's formulation of the supervenience relationship (1998, p. 147).

4.4 Characterising Particulars and Psychophysical Reductionism

One can raise two kinds of question concerning the identity of universals. One at the type level; namely, when is U1 the same universal as U2, and one at the token level; namely, when is one instantiation of a universal identical with another. Clearly, the answer to the latter question is determined by the answer to the former question. For two instantiations of a universal to be identical, they must involve the same universal. Therefore, if properties are universals, the identification of a mental cause and a physical cause, that is the identification of a substance's instantiation of a mental property with its instantiation of a physical property, requires the identification of the mental and physical at the level of types. If, on the other hand, property-instances are CPs, the identification of mental causes with physical causes does not require the identification of mental and physical types, for the question of when two CPs are identical is a question about the identity of a token level entity. This consideration provides the basis for the solution that Robb (1997) offers to the problem of mental causation.³⁸

Robb sets out the problem of mental causation in the following way: Given the acceptance of psychophysical causation, the causal closure principle, and the argument from multiple realisability we are committed to the following three principles:

Relevance: Mental properties are (sometimes) causally relevant to physical events. (p. 180)³⁹

Closure: Every physical event has in its causal history only physical events and physical properties. (p.183)

Distinctness: Mental properties are not physical properties. (p. 182)

Each of these principles is plausible, and yet they appear to be inconsistent. But rather than denying any one of them, Robb argues that if the properties of causation (or in other

³⁸ This proposal is also advanced by Heil and Robb (forthcoming).

³⁹ Robb considers events to be the causal relata, where events are causes in virtue of their properties. Unlike me, Robb considers the causal relata to have non-causal properties, but as he comments, his argument is equally applicable given a more fine-grained theory of the causal relata (1997. p. 181).

words, the properties that events are instantiations of) and hence of mental causation, are CPs rather than universals, the principles can be reconciled.

If events are characterised by universals, and hence universals are the properties in virtue of which events are causes, then given Relevance and Closure, mental universals must be identical with physical universals. Hence the mental and the physical must be identical at the type-level. But this is implausible due to the multiple realisability argument. In support of this one may appeal to the most plausible criterion of identity for universals — that of coextensionality. Whether necessary coextension is sufficient for identity of universals is debatable, but certainly for U1 and U2 to be identical they must apply to the very same particulars. If mental universals are multiply realised by physical universals, mental universals cannot be coextensive with physical universals and hence cannot be identical.

If CPs are the properties in virtue of which events are causes, then to allow Relevance and Closure, mental CPs must be identical with physical CPs. However, the multiple realisability argument cannot be raised against the identification of mental CPs with physical CPs, for it is an argument that arises at the level of types only. The idea that mental properties are multiply realisable — that a mental property can be shared by various physical structures — is incorrect if 'property' is here being read as 'CP', for due to its particularity a mental CP could not be had by different physical structures.

This is not to suggest that the multiple realisability argument does not arise within a CP ontology at the level of types. Hence, take two different kinds of structure s_1 and s_2 . Assuming a resemblance theory of CPs, call the set of CPs that exactly resemble each other in regard to being a pain set M . Both s_1 and s_2 instantiate a CP from set M . Hence, s_1 's pain and s_2 's pain although distinct, exactly resemble each other. Let us say that both s_1 's pain and s_2 's pain are realised by physical CPs, but that the physical CPs do not exactly resemble each other. s_1 's pain is realised by a CP from the set of exactly resembling CPs that is physical₁ (call this P_1), and s_2 's pain is realised by a CP from the set of exactly resembling CPs that is physical₂ (P_2). As P_1 is not identical with P_2 , M cannot be identical with either P_1 or P_2 . Hence mental types are not identical with physical types. Consequently, even if properties are CPs, the identification of mental types with physical types should be avoided, as such an identification faces a modified version of the multiple realisability argument.

But Robb considers himself able to avoid any such problem, for despite identifying mental and physical CPs, he does not identify mental and physical types. That is, he combines a CP monism with a type dualism. However, unlike the type dualist who understands properties to be universals, given Robb's CP ontology the question of whether mental types are epiphenomenal is an illegitimate one, for within a CP ontology types are not entities that make a causal difference.⁴⁰ Consequently, Robb is able to resolve the problem of mental causation by identifying mental and physical CPs hence allowing Relevance and Closure, whilst admitting Distinctness at the level of types hence avoiding the problem of multiple realisation.

Although, for reasons that will be discussed in §4.5 and §4.6, I consider that questions regarding the understanding of a property-instance certainly cannot go ignored within the mental causation debate, I do not consider the shift from universals to CPs to offer a plausible form of psychophysical reductionism. This is because, contrary to Robb, one cannot combine a CP monism with a type dualism. Given a plausible interpretation of types within a CP ontology, CP monism in fact entails type monism. Consequently, if one responds to Relevance and Closure by identifying mental CPs with physical CPs, one must also identify mental types with physical types and thus deny Distinctness. In doing so one will encounter the problem of multiple realisability. On the other hand, if one does not identify mental CPs with physical CPs, one must reject either Relevance or Closure.

Assuming a resemblance theory of CPs, at the type level one has sets of resembling CPs, two CPs being of the same type if and only if they belong to the same resemblance class. If types are interpreted in this way, Robb's position is that mental CPs are identical with physical CPs, but at the type level sets of resembling CPs that are mental-types are not identical with sets of resembling CPs that are physical-types. Now the criterion of identity for classes of resembling CPs depends upon the nature of classes. Two classes are identical if and only if their members are identical. Hence, whether two sets of resembling CPs are identical depends upon whether their members are identical. A mental type is identical with a physical type if and only if it has exactly the same members. Thus given Robb's position, mental CPs are identical with physical CPs, but the set of resembling CPs that is a mental type does not have exactly the same members as any set of resembling CPs that is a physical type. Contrary to this, I shall argue that if mental and physical CPs are identical, as a matter of fact a mental type must have exactly the same

⁴⁰ This is a point to which I shall return in Part 1, §4.6.

members as some physical type. To see why, we need to consider the relation of resemblance.

It is most plausible that the relation of resemblance existing between CPs is an unanalysable primitive relation, for which no eliminative definition is available.⁴¹ However, this is not to suggest that one cannot say anything about the relation of resemblance. Firstly, resemblance is an internal relation. A CP is a member of a certain resemblance class because of what it is. Whether or not two CPs resemble each other wholly depends upon their particular natures. Hence, assuming a CP monism, obviously a mental CP and the physical CP that it is identical to, must enter into exactly the same resemblance classes.

Secondly, resemblance is to be understood as a relation between CPs that admits of degrees, allowing different resemblance classes involving different degrees of resemblance. At the lowest end of the scale, two CPs do not resemble each other at all. At the highest end, two CPs exactly resemble each other. It is only those classes of CPs with the greatest possible degree of unity, that is, sets of exactly resembling CPs, which can be substituted for universals. This can be seen by the formal properties of the relation of resemblance. Whilst all resemblance relations are reflexive and symmetrical, it is only in the case of exact resemblance that the relation of resemblance is transitive. This means that classes of exactly resembling CPs are equivalence classes and thus can be divided into mutually exclusive classes, each providing a substitute for a universal.⁴² Thus returning to the multiple realisability argument, the claim is that a single mental universal is multiply realised by different physical universals. So in terms of CPs, this should be interpreted as the claim that the members of a set of exactly resembling CPs that is a mental type, are realised by physical CPs that do not themselves exactly resemble each other. If one were to argue that one should instead talk about a set of inexactly resembling mental CPs whose members are realised by physical CPs that do not exactly resemble each other, this is to deny the argument from multiple realisability. For sets of inexactly resembling CPs are to be substituted within a theory of universals for different universals,

⁴¹This is because any attempt to define 'resemblance' in terms of something else inevitably leads to a theory of universals rather than one of CPs, because plausible eliminative accounts of resemblance analyse resemblance in terms of identity. For further defence of this point see Campbell (1991, pp. 38-40).

⁴² This point is made by Armstrong (1989, ch. 6).

whose resemblance is most commonly explained in terms of a partial identity.⁴³ And if the mental universals realised by different physical universals are not themselves identical we do not have a problem of multiple realisation in the first place.

Given these considerations we can now consider the problem with Robb's position. Because of Relevance and Closure, Robb advances a CP monism. Hence, let us say that M1 and M2 are two mental CPs, and P1 and P2 are two physical CPs, and that M1 is identical with P1, and M2 is identical with P2. Given the acceptance of the argument from multiple realisability, Distinctness must be interpreted as the claim that the resemblance relation between M1 and M2 is that of exact resemblance, whilst the resemblance relation between P1 and P2 is not. Hence a CP monism combined with a type dualism must therefore allow that although M1 is identical with P1 and M2 is identical with P2, M1 exactly resembles M2 but P1 does not exactly resemble P2. But given the formal properties of the relation of exact resemblance this is false. Obviously as M1 (M2) is identical with P1 (P2) and the relation of resemblance is reflexive, M1 (M2) and P1 (P2) exactly resemble each other. As the relation between M1 and M2 is that of exact resemblance, given the transitivity of exact resemblance if P1 exactly resembles M1, M1 exactly resembles M2, and M2 exactly resembles P2, then P1 must exactly resemble P2. So P1 and P2 must belong to the same exact resemblance class. That is they must be of the same type. CP monism entails type monism.

Of course, this is to assume that there is a single scale of resemblance. If this is not the case, then one may indeed combine a CP monism with a type dualism, for one may argue that although P1 exactly m-resembles P2 (and therefore is a mental type) P1 does not exactly p-resemble P2 (and therefore is not a physical type). Indeed, this appears to be Robb's suggestion. He explains that 'if trope monism is true, a given mental type is a set of physical tropes. But multiple realizability entails that these physical tropes do not themselves resemble one another in the way that members of a *physical* type must: they will be wildly dissimilar physically. So the mental type is not itself a physical type...' (1997, p. 188).

But if the CP theorist can allow that there are various scales of resemblance might the proponent of universals not raise a similar response to the problem of multiple realisability? Take the two universals U1 and U2. Substances s1 and s2 are distinct and

⁴³ For an analysis of the resemblance between universals in terms of partial identity see Armstrong (1989, pp. 103-7).

s1 instantiates U1 and s2 instantiates U2. Now let us say that if two universals are m-identical they are the same mental universal and if two universals are p-identical they are the same physical universal. U1 is m-identical with U2, but U1 is not p-identical with U2. So relative to m-identity, U1 and U2 exemplify a single mental universal, but relative to p-identity they exemplify distinct physical universals. Now this is quite clearly the wrong way to go about things. There are not different kinds of identity for different kinds of things, and there are certainly not different kinds of identity for the same thing. Of course, different criteria of identity for universals may disagree over whether U1 and U2 are identical. Equally, to which type one categorises a property is a relative and often subjective matter. But, assuming that universals have determinate identity, there must be a fact of the matter as to whether U1 and U2 are identical, and if they are identical then one cannot maintain that there is some further identity relation according to which they are not.

And the same is equally true of the relationship of resemblance. There is a fundamental relation of resemblance and hence there is a single fact of the matter about whether the CPs M1 and M2 exactly resemble each other. This, one might respond, is to elevate the status of types within a CP ontology. But just because types are simply sets of CPs that exactly resemble each other and hence are not ontologically additional entities, this does not mean that given a CP ontology there is not a fact of the matter about what types there are. Campbell (1991), goes to great effort to explain that the relation of resemblance is not an arbitrary one. Rather, resemblance is something objectively there in the world. Thus, according to Campbell, 'Of the very many natural kinds, mankind is sensitive to some small proportion, some smaller proportion of which he dignifies with classification and labelling' (1991, p. 31). Indeed, as Armstrong (1989, p. 122) comments there is nothing to stop a CP theory agreeing with a theory of universals about what types there actually are.

Furthermore, if there is no fundamental resemblance relation governing type existence within a CP ontology, we will have problems when it comes to providing a criterion of identity for CPs themselves. How to provide a criterion of identity for CPs will be considered in detail in Part Two, but briefly, for two CPs to be strictly identical they must be the same particular. Assuming a two category ontology, I argue that for two CPs to be strictly identical they must be instantiated by the same substance. This clearly does not provide a sufficient condition for the identity of CPs. One should not, for example, identify a mental and a physical CP merely on the basis that they are instantiated by the same person. What is missing is the further requirement that the relevant properties have exactly resembling natures and hence belong to the same exact resemblance class.

Indeed, regardless of one's understanding of a particular property, this is surely a necessary requirement for any plausible criterion of CP identity. But this entails that if there is to be a single fact of the matter about CP identity (which surely there must be), then there must be a single fact of the matter about whether two CPs exactly resemble each other. That is, there must be a fundamental resemblance relation governing type existence. If there is more than one kind of resemblance relation that determines the existence of types within a CP ontology, then we are left without any plausible criterion of CP identity. Hence, contrary to Robb's suggestion there must be a fundamental resemblance relation governing type existence.

Given that there is a fundamental resemblance relation governing type existence, I would suggest that the only way to combine a CP monism with a type dualism is to allow that CPs have various aspects. Hence, let us say that M1, M2, P1 and P2 are CPs and that M1 is identical with P1 and M2 is identical with P2. P1/ M1 has aspects X1 and Y1, and P2/M2 has aspects X2 and Y2. Furthermore, let us say that aspects X1 and X2 exactly resemble each other, but aspects Y1 and Y2 do not. In virtue of the exact resemblance of aspects X1 and X2, P1/M1 and P2/M2 are the same mental type. In virtue of the fact that Y1 and Y2 do not exactly resemble each other, P1 and P2 are not of the same physical type. By appealing to the notion of property aspects, we have combined a CP monism with a type dualism.

Unfortunately, it is crucial that Robb denies that CPs have any such aspects. It is legitimate to ask of a theory of thick-grained events as the causal relata whether a mental event is a cause in virtue of its mental properties. Therefore, if properties have aspects, then surely it is legitimate to ask whether it is in virtue of a property's mental aspect or its physical aspect, that mental properties make a causal difference in the physical domain. And if this question is indeed legitimate, then Robb has not resolved the problem of mental causation, for in answering this question he is forced to deny (modified versions of) either Closure or Relevance.

Robb, however, does not deem the question legitimate, because he does not consider CPs to have aspects.⁴⁴ Certainly, substances and events have various aspects in virtue of the fact that they instantiate various properties, and a complex of properties has various aspects in virtue of the fact that the complex has various parts. But the 'ball's shape is not roundness in virtue of this or that property it has, it is just roundness full stop.' (Robb

⁴⁴ See Robb (1997, pp.190-1) and Robb (2001, pp. 92-93).

(2001, p. 93)) And similarly a mental CP is not mental in virtue of this or that aspect of the CP. It is just mental full stop. To suggest that CPs have aspects in this sense is, as Robb comments, 'odd'. The problem for Robb is that although it is indeed odd to suggest that properties have aspects in this sense, if there is a fundamental resemblance relation governing type existence, I can see no other way to combine a CP monism with a type dualism.

To summarise, to deny that the relation between the members of a mental type that is multiply realised by physical types is one of exact resemblance is to deny the argument from multiple realisability. But if two mental CPs exactly resemble each other and are identical with physical CPs, they must all have exactly resembling natures. It therefore follows that in actual fact the physical CPs must belong to the same exact resemblance classes and hence be of the same type. One therefore cannot combine a CP monism with a type dualism. So if one accepts the problem of multiple realisation and identifies mental CPs with physical CPs one cannot avoid encountering it. If one attempts to combine a CP monism with a type dualism by denying that there is a fundamental resemblance relation governing type existence, then this leads to serious problems for CP analysis. Alternatively, if one attempts to combine a CP monism with a type dualism by suggesting that CPs have various aspects then, aside from the fact that it is difficult to understand what could be meant by such an 'aspect' of a property, this is to simply relocate the problem of mental causation. An appeal to CPs does not provide a solution to the problem of mental causation in the form a psychophysical reductionism.

4.5 Property Analysis

If Robb is incorrect in thinking that CPs allow one to advance a plausible form of psychophysical reductionism, one may once again raise the question: Does it matter to the mental causation debate whether the properties of causation are universals or CPs? I would suggest that it does. The move from universals to CPs has a number of subtle but very important and wide-ranging effects on the mental causation debate. In order to see what these effects are we need to look ahead to the second and third part of this thesis.

The concern of Part Two of this thesis is with property analysis. Here I shall simply point out three very general reasons why property analysis matters to the mental causation

debate. The interpretation of the premise of psychophysical causation, the independent plausibility of the conclusion of the argument from causal overdetermination (that mental property-instantiations are identical with physical property-instantiations) and the true range of positions within the philosophy of mind with which the premises of the argument from causal overdetermination are compatible, depends upon what the criterion of identity and existence for properties is.

Contrary to the eliminativist and the epiphenomenalist, the premise of psychophysical causation should not be denied. However, it is important to establish what exactly it is that a denial of psychophysical causation truly consists in, and hence what it is that distinguishes eliminativists and epiphenomenalists, from those who would wish to accept this premise. For there to be psychophysical causation, given my understanding of the causal relata as property-instantiations, clearly mental properties must exist and their instantiation must make a causal difference within the physical domain. But to establish whether mental properties exist, one must first consider the criterion of existence for properties. From the fact that a substance falls under a mental predicate, does it follow that there is a mental property that the substance instantiates? If not, how can one tell which properties a substance genuinely instantiates, and hence whether substances instantiate mental properties?

If mental properties do exist, then do they make a causal difference in the physical domain? Certainly, mental statements are appealed to within causal explanations of the physical but what ontological conclusions we should draw on the basis of our causal language again depends upon what a plausible criterion of property identity and existence entails for the correspondence relationship between predicates and properties. The causal statement '*s1* instantiating *p1* caused the instantiation of *p2* by *s2*' is true only if the predicates *p1* and *p2* truly hold of *s1* and *s2* respectively, and this wholly depends upon the properties that the relevant substances instantiate. For the causal statement to be true, it must be the properties of the two substances in virtue of which *p1* and *p2* hold that make the causal difference. For example, the truth-value of the causal statement 'Fred's belief caused him to vote' depends upon whether the properties in virtue of which it is correct to describe Fred as having a belief, are the cause of his voting. If there is psychophysical causation the relevant properties must be mental ones.

The conclusion of the causal overdetermination argument is that mental causes are identical with physical causes. Is such a conclusion independently plausible, or should its

implausibility in fact lead one to question the argument from causal overdetermination? Whether it is plausible that mental causes are identical with physical causes, depends upon whether it is plausible that mental properties are the same properties as physical properties, for in order to identify 's1 instantiating P1 at t1' with 's2 instantiating P2 at t1' one must be able to identify P1 with P2. It is only with a properly worked out criterion of identity for properties that one can establish what must be the case for mental properties to be identical with physical properties, and hence mental causes with physical causes.

Finally, although accepting the three premises of the argument from causal overdetermination, not everyone would accept a psychophysical reductionism. Rather than identifying mental properties with physical properties, an alternative relation between mental and physical properties is advanced which is considered to be compatible with the three premises of the argument from causal overdetermination — mental properties 'supervene upon' or are 'realised by' physical properties. These kinds of physicalist position, although rejecting the identity of mental and physical properties, maintain that mental properties in some way 'depend' upon physical properties. But does it make sense to suggest that properties depend upon one another? And if so, in what sense? What must the relationship between mental and physical properties be in order for there to be a dependence relationship between them? Various relationships have been suggested. For example, it is proposed that mental and physical properties are related as disposition to categorical base or as determinable to determinate. With a criterion of existence and identity for properties, one can establish whether, if, for example, mental predicates are related to physical predicates as determinables to determinates, one should conclude that mental properties are really distinct from physical properties, or whether this in fact entails that mental properties are identical with physical properties or, indeed, that they should be altogether eliminated.

Hence, what the criterion of existence and identity for properties is, is important to the mental causation debate. But in order to provide a criterion of existence and identity for properties, one must first decide upon what properties are — that is, whether they are CPs or universals. Whether one is understanding properties to be universals or CPs, will radically affect the way in which one proceeds in the search for a plausible criterion of existence and identity for properties. For example, the most commonly offered criterion of identity for universals is that of necessary coextensionality. According to it, if U1 and U2 are necessarily instantiated by the very same particulars, then they are the same. However, this criterion of identity has no application if properties are CPs.

4.6 Singularist and Generalist Theories of Causation

There is a further important reason why it matters to the mental causation debate whether the properties of causation are universals or CPs. Part Three of this thesis considers what psychophysical causation might consist in. It discusses the strength of the causal closure principle required in order for one to deny interactive mentalism and asks whether such a principle is plausible. I argue that one's response to this question greatly depends upon what one understands causation to be. In particular, how hard it is to maintain an interactive mentalism depends upon the model of causation that one is basing psychophysical causation on.

But the theory of causation one maintains will in turn be influenced by whether one identifies the properties of causation with CPs or universals. To see why, let us once again consider Robb's position within the mental causation debate. To secure the causal relevance of the mental within the physical domain, given the causal closure of the physical domain, all that Robb considered to matter was the identity of mental and physical tropes. He assumed that given his trope ontology, it did not matter whether mental and physical types were also identical, for within a trope ontology a type dualism does not have to face the issue of type epiphenomenalism. This is because within a trope ontology, types are not entities that make a causal difference. But I would suggest that it is only the case that Robb's type dualism does not have to face the issue of type epiphenomenalism, if Robb is assuming a singularist theory of causation. That is, it is only true that types do not make a causal difference, and hence that Robb can ignore the question of whether the mental and physical are identical at the type level, if one assumes a singularist theory of causation.

Returning to the distinction that I made between generalist and singularist theories of causation, a singularist denies that singular causation is grounded by type level relations, whilst generalists maintain that for particular states to be causally related at the token level, they must instantiate types of states that are suitably related. Most plausibly, laws furnish causation's generalist component. Consequently, the causal relata must exemplify types such that these types are lawfully connected. Singularist theories disagree — a causal relation may fail to be subsumable under any law, indeed causal laws may be grounded in singular causal relations rather than the other way around.

Robb's position is only plausible if he is assuming a singularist theory of causation, for even if property-instantiations involve CPs, assuming a generalist theory of causation, types still ground causation. So given a generalist theory of causation if, contrary to fact, one could combine a trope monism with a type dualism (as Robb suggests) one would still not have resolved the problem of mental causation. Mental causes would be identical with physical causes because mental tropes would be physical tropes, but it would be wholly legitimate to ask whether the mental cause was a cause in virtue of being a mental type or a physical type. If it were in virtue of being a physical type, then mental types would be epiphenomenal. If it was in virtue of being a mental type, then how could one reconcile this with the causal closure principle? The fact that Robb assumes that his position cannot legitimately be accused of type epiphenomenalism means that he is assuming a singularist theory of causation.

And this assumption is entirely justified, for CPs are most plausibly incorporated within a singularist theory of causation. At the most basic level, to base CPs within a generalist theory of causation seems like the wrong way to go about things. It would mean that a CP makes the causal difference that it does because of the type level relations that it falls into. That is, a CP makes a causal difference because it is a member of a certain resemblance class, which is related to another resemblance class in the relevant way. But surely the causal differences that a CP makes depends wholly upon it, upon its nature. Surely, it does not depend upon whether certain relations exist between the sets that it is a member of.

Furthermore, the kind of nomological theory of causation that a CP ontology is consistent with is an implausible one. If one considers singular causation to be grounded by laws, then the plausibility of one's theory of causation will depend upon the plausibility of the account one gives of a law. A generalist theory of causation that analyses laws of nature in terms of regularities is as implausible as the theory of laws upon which it is based. And the problem with maintaining a generalist theory of causation, if the properties of causation are CPs, is that within an ontology of CPs it is difficult to advance anything other than a regularity theory of laws.

If properties are universals, one can escape the regularity theory of laws and advance a far stronger account of laws. Thus according to Armstrong (1997), laws of nature express relations of necessitation between universals. If universals U1 and U2 are lawfully related,

and s1 instantiates U1, then under the appropriate circumstances, s1 must also instantiate U2. If, on another occasion, U1 is instantiated by a different substance in similar circumstances, we can understand why it results in the same effects, because we are dealing with the very same universal. So we can explain why substances that instantiate the same universal in similar circumstances enter into the same nomological relations.

However, it is arguable that with only the apparatus of CPs, we cannot go beyond a regularity theory of laws.⁴⁵ As we have seen, despite the fact that a CP is not a unifying entity, one can explain types without universals by appealing to sets of resembling CPs. In particular, sets of exactly resembling CPs can be made to mimic universals. It might therefore be thought that they could play a nomic role as well as universals. If two CPs exactly resemble each other, then in similar circumstances, they will enter into the same nomological relations. But the problem is that of explaining why, when they are in similar circumstances, exactly resembling CPs enter into the same nomological relations; that is, why exactly resembling CPs display regular behaviour. To simply respond that if two CPs exactly resemble, then they just do display this regularity, is not to go beyond a regularity theory of laws. Hence, given only CPs, it is difficult to go beyond a regularity theory of laws, and thus advance a plausible nomological theory of causation. That is, it is implausible to advance a generalist theory of causation, if the properties of causation are CPs.

To clarify this point, attention can be drawn to a narrower understanding of generalism. According to it, whether or not two particular states are causally connected depends upon things that happen elsewhere or elsewhen — causation is not a purely local matter. Hence, a naïve regularity theory of causation is, in this sense, generalist. The naïve regularity theorist maintains that to say that c causes e is simply to say that e succeeded c and that events of type c are regularly and indeed always succeeded by events of type e. Hence according to it, a particular sequence is causal by virtue of being an instance of a general pattern. The same particular sequence coupled with no general pattern would not be causal. Less obviously, Lewis' Counterfactual Theory of causation (1973) is also generalist in this sense. To assess whether 'e causally depends upon c' is true one must evaluate the counterfactual: 'If c were not to occur, then e would not occur.' Lewis maintains that this counterfactual is true if and only if possible worlds in which c and e fail to occur, are more similar to the actual world, than possible worlds in which c fails to occur, but e does occur. But Lewis considers interworld similarity to be in part determined

⁴⁵ This is the main reason why Armstrong prefers universals to tropes. See Armstrong (1997, pp. 23-4).

by nomological similarities across these possible worlds, and maintains a sophisticated version of the regularity theory of laws.

Contrary to this, one might consider causation to be a purely local matter. Thus, for example, a transference theory of causation is singularist in this sense. Transference theories maintain that 'c causes e' if and only if it is true that something is transferred from c to e. These theories are singularist because according to them whether or not two entities are causally connected simply depends upon whether there is a transference of a quantity between them.⁴⁶ Whether there is such a transfer does not depend upon what is happening elsewhere or elsewhen. (They are also singularist in the first sense, because they do not consider causation to be grounded by nomological relations ⁴⁷).

Now, unlike the naïve regularity theory and Lewis' counterfactual theory, which are generalist in both senses, Armstrong's theory of causation is only generalist in the first sense. According to Armstrong, causation is grounded by laws. Hence, Armstrong is appealing to a type level relation in order to ground causation. But it is not generalist in the second sense, because causation is a purely local matter — whether or not two states of affairs are lawfully connected, does not depend upon what is happening elsewhere or elsewhen — it depends entirely on the particular states of affairs involved in the causal relation.⁴⁸

If, however, one denies that properties are universals, then this is clearly not an option. Because the CP ontologist cannot get away from a regularity theory of laws, any generalist theory of causation that is available to the CP ontologist, must be generalist in both senses. That is, if the CP ontologist adopts a generalist theory of causation, this commits him to the idea that causation depends upon what happens elsewhere or elsewhen. Hence, if the CP ontologist was to advance a generalist theory of causation, then whether or not a CP makes a causal difference would depend entirely upon whether

⁴⁶ They are reductive singularist positions, because singular causation is wholly analysable in terms of the quantity that is transferred from cause to effect.

⁴⁷ Hence, for example, as Ehring (1997, pp. 44-5) argues, although energy/ momentum transfers are law-governed, the energy transference theory is not a generalist theory, because this requirement is not part of this theory of causation.

⁴⁸ The distinction between singularist and generalist theories of causation that I make is similar to the one that Ehring (1997, ch. 1) makes. However, Ehring does not distinguish between narrow and broad forms of generalism. In conflating the two kinds of generalism he does not provide a place for Armstrong's theory of causation.

it was a member of an exact resemblance class with which other exact resemblance classes form the right pattern. But surely, whether or not a CP makes a causal difference does not depend upon whether it is an instance of a certain pattern — it does not depend upon what is happening elsewhere or elsewhen. It depends entirely upon the nature of the relevant CP. Hence, if the properties of causation are CPs, one must turn to singularist theories of causation.

Thus, although, contrary to Robb, I do not consider that one's understanding of the properties of causation directly affects the plausibility of positions within the mental causation debate, I do consider one's understanding of the nature of properties to have a general, but very important influence, on the way in which one proceeds within the mental causation debate. For firstly, as I shall argue in Part Two, what the criterion of existence and identity for properties is, is crucial to the mental causation debate, but obviously to provide a criterion of existence and identity for properties one must first decide upon what properties are. Secondly, in Part Three of this thesis, which discusses the plausibility of various strengths of causal closure principle, a central claim is that their plausibility depends upon the theory of causation that is being assumed. But the theory of causation that one adopts will be heavily influenced by what one understands the properties of causation to be.

5

Why the Properties of Causation are Characterising Particulars

5.1 States of Affairs

Having considered the reasons why it matters to the mental causation debate what the properties of causation are, I will now discuss the reason why I think it is most plausible that the properties of causation are CPs. Because the doctrine that properties are universals, has, as Heil and Robb (forthcoming) maintain, a status within the philosophy of mind such that it deserves to be called a dogma, I will explain in reasonable depth why I consider this 'dogma' to be incorrect. Put simply, my argument is that the causal relata are not states of affairs, and that if the causal relata are not states of affairs then the properties of causation must be CPs. I would not wish to claim originality for the individual arguments for the existence of CPs that I put together within this section. Most have already been made by other metaphysicians. Chief among these are C. B. Martin and E. J. Lowe who both go to great lengths to defend the existence of properties as particulars, within the context of a substance ontology.

One interpretation of a property-instantiation is as a fact. 'The apple instantiating the property of redness at time t_1 ' is to be understood as the fact that the apple is red at time t_1 . One can distinguish between three non-equivalent understandings of a fact. According to the first two propositional conceptions, facts are not part of the world, but are descriptions of it; Ducasse (1940) and Carnap (1947) maintain that a fact is a true proposition, and Moore (1953) considers a fact to be the truth of a proposition. According to the third position, facts or 'states of affairs' are not merely descriptive of the world, but are part of it. States of affairs are not true propositions; they underlie true propositions. Hence, according to Armstrong 'A state of affairs exists if and only if a particular ...has a

property or, instead, a relation holds between two or more particulars.' (1997, p.1)⁴⁹. Thus, for Armstrong, the instantiation of a property by a substance, that is, a property-instantiation, is a state of affairs. For example, 'the redness of this dress' refers to a state of affairs, 'the dress being red', which includes a substance (the dress) and a universal (redness) as constituents. While the first two interpretations of facts treat facts as wholly abstract entities, states of affairs are complexes which have concrete constituents and are therefore themselves concrete. To this extent, states of affairs are suitable candidates for the causal relata.

But there are significant problems with admitting states of affairs as entities, let alone as the causal relata. In the first place, it is difficult to provide plausible existence and identity conditions for them. In particular, the 'Slingshot Argument' claims to reduce all facts to a single fact.⁵⁰ In itself, this provides a strong enough reason to resist the claim that the causal relata are facts, but equally worrying is the ontological categories from which states of affairs are constructed. I shall argue that whether one admits Armstrong's states of affairs as ontologically additional entities rests entirely upon whether one accepts Armstrong's notion of a substance. Armstrong understands substances to be 'thin particulars', that is, particularising entities which themselves have no properties. This notion of a substance is deeply problematical, but it is one that Armstrong's states of affairs ontology cannot be detached from. This provides us with a further very strong reason to resist the ontology of states of affairs. But this is to get ahead of myself. Here let it simply be stated that my premise is that if one can, one should avoid admitting Armstrong's states of affairs as the causal relata.

However, if the arguments that Armstrong presents in *A World of States of Affairs* are correct, then it is in fact implausible to interpret a property-instantiation as anything other than a state of affairs. Furthermore, according to Armstrong, this is the case regardless of whether property-instances are CPs or exemplifications of universals. In the following, I argue that this is incorrect. One can plausibly avoid identifying property-instantiations with states of affairs. However, in order to do so, one must identify the properties of causation with CPs.

⁴⁹ Also see Mellor (1991) who also maintains that facts are ontologically basic, out of which both universals and particulars are abstractions.

⁵⁰ For a version of the slingshot argument see Davidson (1984c). The slingshot argument is critically examined by Olson (1987) and Neale (1995)

To show how crucial Armstrong's understanding of a substance is to his state of affairs ontology, in §5.2, I begin by considering Kimean events, which are commonly accused of being nothing other than facts. Despite the many similarities between Kim's events and Armstrong's states of affairs, in §5.3 I argue that whether Kimean events really are states of affairs wholly depends upon Kim's understanding of a 'substance.' It is only if the substances within Kim's account are thin particulars that Kimean events are states of affairs. If substances are thin particulars then property-instantiations must be states of affairs. Furthermore, this is true regardless of whether the properties involved are universals or CPs. However, there are two further understandings of a substance (Armstrong rejects the first, but fails to recognise the second). Neither requires one to identify a property-instantiation with a state of affairs. The first reduces substances to bundles of properties. Within such a one-category ontology, as Armstrong himself recognises, one need not identify property-instantiations with states of affairs. This one-category ontology can only be given any plausibility if properties are CPs. More plausibly, one can assume a two-category ontology, but not identify substances with thin particulars. A substance ontologist admits the ontological category of substance as fundamental, yet does not identify substances with thin particulars. Within such an ontology there is no need to admit states of affairs as ontologically additional entities. But in §5.4, I argue that given this ontology, if property-instantiations are the causal relata, one cannot maintain a neutral stance on whether properties are universals or CPs — they must be CPs. Hence, either the causal relata are states of affairs which requires one to admit the implausible ontological category of thin particular, or the properties of causation are CPs. These three different options are set out in table 1 below.

<u>Substances</u>	<u>States of Affairs?</u>	<u>Properties of causation must be CPs?</u>
Thin Particulars	Yes	No
Reduced to bundles of properties.	No	Yes
Substance ontology	No	Yes

5.2 Kimean Events

Given that the causal relata are property-instantiations, one of the most obvious candidates for the causal relata are Kimean events. According to Kim: '.....each individual event has three unique constituents: 'a substance (the 'constitutive object' of the event) a property it exemplifies (the 'constitutive property' or 'generic event'), and a time' (1993c, p. 35).⁵¹ Event (s1, P1, t1) exists just in case the substance s1 exemplifies the property P1 at time t1. Kim also offers a criterion of event identity: Events (s1, P1, t1) and (s2, P2, t2) are identical if and only if s1 = s2, P1 = P2, and t1= t2 (1993c, p. 35).

Although the term 'universal' is referred to only once or twice within Kim's writing on events, that properties are universals is clearly an underlying premise throughout his discussion. For example, Kim assumes that talk about a kind of event is equivalent to talk about a property: 'Every event has a unique constitutive property...And, for us, these constitutive properties of events are generic events' (1993b, p.12). As it is here being assumed that a property is a unifying entity, properties must be universals. Hence according to Kim, for a substance to exemplify a property at a time is for a substance to exemplify a universal at a time. By a 'substance' Kim means '...things like tables, chairs, atoms, living creatures, bits of stuff like water and bronze' and according to him 'there is no need to associate this notion with a particular philosophical doctrine about substance' (1993c, p.33). But presumably, given the great implausibility of a bundle of universals theory, one can assume that Kim considers substances to exist over and above properties, hence maintaining a two-category ontology.

How should one interpret Kimean events? More specifically, are Kimean events just states of affairs? Kimean events do not include as simple events occurrences such as carnivals and disasters, because Kim purposefully underplays the difference between events and states. If one preserves the term 'event' for an entity that possesses a temporal dimension as well as a spatial one, hence identifying events with changes in the properties of a substance, Kimean events should be classified as states.⁵² The difficult question is whether Kimean events are states of substances or states of affairs? That is, given Kim's understanding of property-exemplifications should, for example, the apple exemplifying

⁵¹ For Kim's theory of events see Kim (1993b) and (1993c). Another form of property-exemplification account is advanced by Goldman (1970).

⁵² Or more properly, given the distinction between an event and an aspect of an event, a property-exemplification is an aspect of a state

the property of redness at time *t* be understood as the state; the apple being red at time *t*, or the fact that the apple is red at time *t*?

At first glance, Kim's events appear very similar, if not identical with, Armstrong's states of affairs. According to Kim, if the table exemplifies the universal property of brownness at time *t* there is an event, whilst according to Armstrong if the table instantiates the universal property of brownness there is a state of affairs. Both have properties, where properties are universals. Both appeal to a further ontological category, a 'particular' or 'substance'. And both assume there to be an event/ state of affairs if there is some sort of relationship between this substance and the property; Armstrong talks about 'instantiation', but he acknowledges that one could equally talk about 'exemplification' (1997, p. 115). Furthermore, their criteria of identity for these entities are remarkably similar — both maintain that the identities between the respective entities depend upon their internal constitution and structure; if they contain the same constituents and these are organised in the same way then the entities are identical.⁵³

It would therefore appear that Kim is confusing events with facts, or more properly states of substances with states of affairs. But are the many similarities merely apparent? This in fact wholly depends upon whether the 'substances' of Kimean events are to be identified with the 'particulars' of Armstrong's states of affairs. Contrary to Kim's suggestion, one needs to know something about the ontological nature of substances within his account, in order to assess its plausibility.

5.3 The Properties of Causation and Three Different Understandings of 'Substance'

i) Substances as thin particulars

What is this 'substance' that exemplifies properties? One's immediate response may be to identify the substance with a concrete object. It is, for example, the apple conceived as a concrete object that exemplifies the property of redness. But arguably this relationship between a substance and its properties is misconceived. It is not the apple, conceived as

⁵³ For Armstrong's criterion of identity for states of affairs see Armstrong (1997, pp. 131-2).

a concrete object, which has the property of redness. The redness is a constituent of the apple, and there is a further constituent of the apple, the thin particular, and it is this, which exemplifies the redness. The thin particular possesses the property of redness but is distinct from this property and any other property it may possess, in itself only differing from other thin particulars numerically. It is, furthermore, a particularising entity; it is because a set of properties are possessed by it that one has only one substance. Hence, one may say that it is the apple that exemplifies the property of redness, but in doing so one is understanding the apple as a thin particular — as something which in itself has no properties.

Therefore, one understanding of a Kimean event is as an exemplification of a universal by a thin particular at a time. If one interprets Kimean event in this way, then they will contain exactly the same constituents as Armstrong's states of affairs, because the substances within his states of affairs are also thin particulars. Furthermore, if Armstrong's argument for the existence of states of affairs as ontologically additional entities is correct, and one interprets property-exemplifications in this way (and one is ontologically serious) then Kimean events must be states of affairs.

According to Armstrong, unless one wishes to 'abandon ontological seriousness' one should accept the truthmaker principle — the requirement that all truths have a truthmaker (1997, p. 135). But given this principle, if one accepts the existence of thin particulars and properties, one must accept monadic states of affairs as an ontologically additional category. A thin particular St1 instantiates a property P1. But according to the truthmaker principle there must be something in the world that serves as the ontological ground for this relation — something is needed to 'weld' substance and attribute together (1997, §8.11). This can be neither St1 nor P1 considered individually, nor can it be the pair St1 and P1, for just because both exist does not mean that there need be an instantiation relationship between them. According to Armstrong, the truthmaker must be the state of affairs 'St1 instantiating P1', for it is the state of affairs which brings the substance and property together in a non-mereological form of composition. Hence, if the substances to which the property-exemplificationist account refers are thin particulars, and thus one accepts the substance-attribute distinction, then states of affairs must be taken ontologically seriously and 'St1 exemplifying P1 at t1' must be a state of affairs. This argument in fact stands regardless of whether the properties that thin particulars instantiate are universals or CPs. If a thin particular instantiates a CP, something

ontologically additional is still required to weld the two ontological categories together. (See Armstrong (1997, § 8.12)).

ii) Substances as reducible to bundles of properties

These considerations also allow one to see why within a one-category ontology one need not identify property-instantiations with states of affairs. Given a bundle theory, as there is no irreducible substance, states of affairs are not required to weld properties together with anything — the property instantiation just is a property from a bundle of properties.⁵⁴

It also enables one to see why some philosophers associate states of affairs with universals rather than CPs.⁵⁵ The question of whether substances are ontologically additional entities and the question of whether properties are universals or CPs are not independent. If the only characterising entities are universals, then there must be some kind of ontologically additional substance, for a bundle of universals theory leads to the undesirable consequence that Leibniz' principle of identity of indiscernibles is a necessary truth. Hence, a two-category ontology, and thus according to Armstrong, states of affairs, naturally follow if properties are universals. On the other hand, those who identify properties with CPs, commonly incorporate CPs within a bundle theory.⁵⁶ Certainly, a bundle of CPs theory is more plausible than a bundle of universals theory, because the bundles will be particulars. Furthermore, it is commonly argued that if properties are particulars, then there is no need to postulate the additional category of substance. The role of a substance is to provide an object with particularity, and if properties can play this role, then surely substances are not needed.⁵⁷ Hence, given a desire for ontological economy one should incorporate CPs within a one-category ontology. And given a one-category ontology, as there is no substance, and thus no ontological tie, states of affairs are not required as ontologically additional entities. If this reasoning were correct then one would be faced with a choice. Either maintain a two-category ontology, in which case

⁵⁴ Although note that a bundle theory does not automatically avoid states of affairs, for given Armstrong's account, states of affairs are still required to weld constituents of a bundle together. See Armstrong (1997, p. 97).

⁵⁵ For example, Ehring (1997, p. 11) assumes that a state of affairs is an instantiation of a universal by any object, and that CPs are to be contrasted with states of affairs.

⁵⁶ For example, Williams (1966) and Campbell (1991) who both advance a bundle of tropes theory.

⁵⁷ See, for example, Campbell (1991, p. 21) who advances claims such as these to defend a bundle of tropes theory.

one must admit states of affairs as ontologically additional entities, or maintain a bundle of CPs theory. In other words, either states of affairs are the causal relata, or the properties of causation are CPs.

Given the choice between a bundle theory and a states of affairs ontology, with Armstrong one might prefer to opt for states of affairs. Even if properties are CPs, it is implausible to incorporate them within a bundle theory. Ontological economy certainly should not be the sole driving force behind one's ontology, and as Armstrong argues, regardless of one's understanding of a property, one should not substantialise them, because they are not 'suitable to be the ultimate constituents of reality'.⁵⁸ Indeed, if one understands properties to be ways, one will consider those that maintain that properties are capable of independent existence to be making something of a category mistake. Properties are not substances, but ways that substances are. It is therefore inconceivable that they could exist independently of substances. The rejection of a bundle theory, Armstrong argues, leads us to accept states of affairs.

Does the rejection of a bundle theory also lead to the rejection of CPs? In support of this it might be asked why, if substances are particularising entities, do we also need properties to be particulars as well. The particularity of a CP would appear to be rendered redundant within a two-category ontology. In response to this, I would suggest that it depends upon what one means by a 'substance'. I am not really bothered whether properties are universals or CPs within an ontology that includes thin particulars, because as will become clear, I find the notion of a thin particular wholly implausible. But, contrary to Armstrong, I do not think that this requires me to accept a bundle theory. There is in fact a far more plausible understanding of 'substance' as an ontologically additional entity, which does not require one to admit states of affairs as ontologically additional entities. Importantly, it does require one to admit CPs. The most plausible way of incorporating CPs into this ontology is to maintain that the particularity of a CP is dependent upon the particularity of the substance that instantiates it.⁵⁹ This is not to suggest that the particularity of the CP just is the particularity of the substance, for this would be to collapse CPs into instantiations of universals. Rather, the particularity of a CP is dependent upon, but distinct from the particularity of a substance. It is this alternative understanding of a substance to which I now turn.

⁵⁸ For a discussion of these points see Armstrong (1997, p. 99) and Armstrong (1989, p. 114-5).

⁵⁹ This is the position of Martin (1980).

iii. Substance ontology

The bundle theories' resistance to a two-category ontology seems to come down to the fact that they find the idea of a 'thin particular' unattractive.⁶⁰ As Armstrong himself acknowledges, the notion of a thin particular is notoriously elusive and indeed dubious — for how can an entity lack properties? The fact that Armstrong understands the relation between a state of affairs and its constituents in a way reminiscent of the Fregean notion of the unsaturatedness of concepts goes some way to removing the mystery which surrounds the notion of a thin particular. (1997, p.29) The state of affairs type is unsaturated, if it 'has one or more blanks as part of its nature' and properties and thin particulars fill these blanks. (1997, p. 29) A thin particular is what we have left when by an act of 'partial consideration' we consider the state of affairs without its properties. Hence, it comes as no surprise that we find the notion of a thin particular strange when we try to consider it in itself, for it is something which only exists within a state of affairs. However, this does not remove the fact that thin particulars are epistemologically suspect, for having no properties, they cannot be experienced. And even if it is legitimate to respond to this by appealing to our experience of the particularity of the thin particular, ontological problems remain. For however intimately connected thin particulars and properties are, the fact remains that Armstrong wishes to admit thin particulars as a distinct ontological category. Hence, he must attempt to answer such questions as: How if thin particulars have no properties, can they be distinguishable from one another, and thus provide the particularising element? And how much of the whole object is the thin particular?

Armstrong's states of affairs do, however, avoid the threat of a Bradleian Regress that two-category ontologies face. Returning to Kimean events, Kim maintains that substances 'exemplify' properties but he never adequately clarifies what he means by this relationship. What Kim cannot do is identify the relationship of 'exemplification' with an ontological relation. If a Kimean event consisted in a thin particular, properties and in addition an ontological relation of exemplification tying the two together, then one would be faced with a Bradleian regress, for in order to put these things together, one must assume some further relation which ties these three constituents together. But then if a further relation is posited, what it is that ties these together — another relation will be needed — ad infinitum. Kim can avoid this problem in the same way that Armstrong (1997, §8.12) arguably does. Armstrong does not consider the relationship of instantiation existing between thin particular and universal to be an ontologically additional component to states

⁶⁰ For example, see Campbell (1991, § 1.4).

of affairs. Rather, it is a state of affairs that holds the constituents of property and thin particular together, and the relationship of instantiation is nothing over and above this state of affairs.

However, an alternative response would have been to appeal to the plausible idea mentioned earlier, that properties are ways that substances are. Armstrong is himself keen to promote properties as ways precisely because it serves to remove any temptation to substantialise them (via a bundle theory). However, if properties are really ways that substances are, then this in itself removes the threat of a Bradleian regress. Properties and substances are so intimately connected that there is no need to admit anything in addition to property and substance to hold them together. Armstrong cannot take this route, for it is precisely because something extra is needed to glue substance and property together, namely a state of affairs, that states of affairs must be admitted as ontologically additional entities. Given his commitment to states of affairs, it is therefore questionable that Armstrong can take the claim that properties are ways seriously. But it does introduce an alternative way of understanding property-instantiations as states of substances, rather than states of affairs, which does not require one to admit the ontologically dubious category of thin particular and which does allow one to take seriously the plausible claim that properties are ways.

Armstrong fails to consider a substance ontology of the kind maintained by Lowe (2001b, forthcoming a, forthcoming b). Like Armstrong, Lowe admits the ontological categories of substance and property. And like Armstrong, Lowe would reject any proposal that properties could be independent entities. Indeed, as Lowe (forthcoming b, pp. 17-20) argues, unlike Armstrong's position, a substance ontology is able to take the claim that properties are ways seriously, for properties are entities that ontologically depend upon substances (while the latter do not ontologically depend upon anything). However, a substance ontology rejects the problematical notion of a thin particular. According to it, a substance is not something that is wholly separate from properties in the sense that a thin particular is; and for this reason, states of affairs do not weld substance and properties together. But nor should it therefore be concluded that a substance just is a state of affairs, for unlike states of affairs, substances are not complexes (forthcoming a, p. 8). That is, the relationship between properties and substance is not that of a part to a whole. Rather, according to Lowe substances comprise a basic, irreducible category, which do not have constituents, but upon which properties ontologically depend.

If a property-instantiation is fitted into this third understanding of a substance, it is not to be identified with a state of affairs. That is, 'The substance instantiating the property p at time t1' is not to be identified with a state of affairs. This is because it is not a complex in which the categories of substance and property are parts.

This would suggest that whether Kimean events are in fact states of affairs depends upon whether the relationship between a Kimean event and the ontological categories to which it is related (that of substance and property) is to be understood as one of a whole to its parts. The fact that Kim considers substances and properties to be ontological constituents of an event suggests it is nothing other than a theory of states of affairs.⁶¹ But the crucial point is that a property-instantiation need not be interpreted in this way within a two-category ontology.

5.4 Are the Properties of Causation In Re Universals or Characterising Particulars?

Given a substance ontology, it is incorrect to maintain that a substance is related to its properties as a whole is to its parts, regardless of one's understanding of a property. Hence, within such an ontology, regardless of whether one understands properties to be CPs or universals, property-instantiations are not states of affairs. Even so, the claim that the properties of causation are universals should still be resisted. This is because unless the causal relata are states of affairs, the properties of causation must be concrete entities, but to understand universals as concrete entities is implausible, for it requires one to adopt an in re understanding of universals and thus face the problem of exactly how a single in re universal can be shared by more than one thing at the same time.

Unless causation is a relation between states of affairs, the properties of causation must be concrete. As discussed in Part One, §1.2, if there are causal relations between non-concrete entities, their non-concreteness must not derive from the nature of the causal relata itself, but the non-concreteness of the object that in this particular case the causal relation is between. To this extent Armstrong's states of affairs are suitable candidates for the causal relata, as are states of objects. However, in order for Armstrong's states of

⁶¹ See, for example, Kim (1993c, p. 35).

affairs to be concrete it is not necessary that the property that the thin particular exemplifies is concrete. This is precisely because Armstrong's states of affairs are structured complexes. A state of affairs is 'non-mereologically' composed by the ontological categories of substance and property. Therefore, a concrete state of affairs must have at least one constituent that is concrete. It is not required that all of its constituents are concrete, for as long as part of it is spatio-temporally located, it too is spatio-temporally located. Hence, for a state of affairs to be concrete it does not follow that the properties that partly constitute the state of affairs must themselves be concrete, provided that a concrete substance also constitutes the state of affairs.⁶² Thus, for example, the state of affairs that is the apple instantiating the universal redness is concrete, because the apple (a thin particular) is located in space and time. This is the case even if the ontological category of universals is abstract and hence the property of redness which the apple instantiates is abstract. In short, if causation is a relation between concrete entities, and the causal relata are states of affairs, then the question of whether the properties of causation are abstract or concrete entities can be left open.

The same is not true of a concrete state. If 'a substance instantiating a property' is to be understood as a state of an object rather than a state of affairs, then the relationship between the property-instantiation on the one hand, and the property and substance on the other, must not be that a whole to its parts. Unlike concrete states of affairs, which are constituted by things which possess spatio-temporal locations, concrete states do not contain things but are rather themselves the things that are spatio-temporally located. Hence it follows that the properties of a concrete state must themselves be concrete. More specifically, if a property-instantiation is incorporated within a substance ontology, and hence properties are simply ways that substances are, it is difficult to understand how a substance might instantiate a property from a realm that is altogether different from it, whilst preserving the intimate relationship that supposedly exists between substance and property. Indeed, if properties are ways that substances are, then surely properties must be located where the substance is. Therefore, the properties that a concrete substance instantiates must themselves be located in space and time.

CPs of concrete objects are themselves concrete; being located in the substances that instantiate them. The red of this apple, for example, has a definite location; it occupies the same place as the apple. More questionable is whether a universal instantiated by a

⁶² Although, I find it difficult to understand how a concrete substance, even a thin particular, could exemplify a non-concrete property.

concrete object is itself concrete. For universals to be concrete entities it is natural to assume (although it is not entailed by) an Aristotelian conception of universals; universals only exist in their individual instances — there are no uninstantiated universals. A location for universals is then found by locating them wherever their instances are, for presumably if universals exist in space-time, they exist where they are instantiated. To combine a theory of properties as universals with this in re assumption is to understand a 'universal' in the sense advanced by the immanent realist. According to this position the universal is 'wholly present' in each of its instantiations, or more specifically, in each of the substances which instantiates it. Thus, for example, my dress and this apple are identical in their colour, because the universal 'redness' is wholly present in both entities. From this in re conception of universals, given that one assumes that all substances are in space and time, it therefore follows that all of the universals exemplified by substances are in space and time.

But immanent realism is implausible. According to the immanent realist, a universal is present wherever the substance that instantiates it is. As a universal can be multiply instantiated by non-overlapping substances, this means that a universal can be present in different places at the same time. The universal redness is, for example, present both in my dress and in this wine. The immanent realist thus has to admit that there are things that can exist in different places at the same time.

Of course, he would not be forced to such a conclusion if his claim was simply that each substance that instantiates the same universal instantiates a different part of the same universal. But this either reduces property-instances to CPs or does nothing to resolve the problem, for if a substance instantiates only a part of a universal, then this part itself either is or is not a particular. If it is a particular, then one is in fact understanding property-instances as CPs, not instantiations of universals. Alternatively, if each of the parts of a universal which the substance instantiates is itself multiply instantiated, then one is still committed to admitting that one thing can have more than one location at the same time — the only difference being that the thing that is multiply located is a part of a universal as opposed to a complex universal itself. Hence, the immanent realist must maintain that a universal is wholly present in each of the substances that it is instantiated by, and therefore that a universal can be wholly present in more than one place at a time. But how can this be so? If a concrete entity, say a table, is wholly present in a place at a time t_1 , it therefore follows that it is not in any other place at t_1 . How can a universal be any different? The immanent realist's typical response is that although it is implausible that

particulars like tables and chairs can be wholly present in more than one place at the same time, universals can be multiply-located precisely because they are not such entities and hence do not follow the same rules. But unless the immanent realist provides us with some indication of how these rules can be plausibly formulated there is no reason why this claim should be accepted.⁶³ For this reason, I would reject the proposal that universals are concrete entities.

CPs clearly do not face this problem. Unlike in re universals, given an in re conception of CPs, a property is not only 'wholly present' in the substance that instantiates it, but is incapable of existing elsewhere. Hence, one is not required to accept the problematical notion of one thing being wholly present in more than one place at the same time. Although an in re conception of CPs and an in re conception of universals are similar in so many ways, their one point of difference — that a CP cannot be multiply instantiated — must lead one to accept a theory of CPs and reject an in re conception of universals. Hence, although a theory of universals need not lead to states of affairs, the kind of universal that has to be assumed in order to avoid them is metaphysically dubious.

To summarise the arguments of this section, whether or not one admits states of affairs depends upon one's understanding of a substance. If substances are thin particulars, then regardless of whether the properties that substances instantiate are universals or CPs this leads to a state of affairs ontology. However, thin particulars are ontologically dubious. Given the denial of thin particulars, substances are either reduced to bundles of properties or understood in the context of a substance ontology such as Lowe's. Neither admits states of affairs as ontologically additional entities. If one maintains a bundle theory then it is most plausible that properties, and hence the properties of causation, are CPs — otherwise Leibniz' principle of identity of indiscernibles must be a necessary truth. If one maintains a substance ontology, then it is most plausible that properties, and hence the properties of causation are CPs. This is because the properties that such a concrete substance instantiates must themselves be located in space and time. But it is only plausible to suggest that universals are concrete entities if one assumes an immanent realism. However, immanent realism is implausible for it requires universals to be wholly present in more than one location at the same time. The same problem clearly does not arise with regard to CPs. Given an in re conception of CPs, a CP is wholly present in the

⁶³ That they cannot be plausibly formulated is suggested by Lowe (forthcoming a, pp. 5-6, forthcoming b), Heil and Robb (forthcoming, pp. 20-21), Oliver (1996).

substance that instantiates it, but equally it is incapable of existing elsewhere. For these reasons, I would suggest that it is most plausible that the properties of causation are CPs.

The Homogeneity of the Causal Relata

Before leaving the discussion of the causal relata, I want to consider one final way in which the causal relata might be thought to play a central role within the mental causation debate. The interactive mentalist can respond to the problem of causal overdetermination by arguing that one can consistently accept the premises of the argument from causal overdetermination whilst denying its conclusion. One way of doing this is to argue that the conclusion that mental causes are identical with physical causes only follows if one is assuming the homogeneity of the causal relata. Lowe's dualist response to the problem of causal overdetermination is based upon a denial of the homogeneity of the causal relata. According to Lowe, mental and physical causation are causation by different kinds of entity.

Lowe (1996, 1999a, 2000) maintains that the self is a simple substance that is distinct from the body or any substantial part of it. Although maintaining an interactive mentalism, he rejects the proposal that the self causes bodily behaviour by initiating a particular physical event. According to most forms of interactive mentalism, psychophysical interactionism is explained in the following way: A mental event (for example, the desire to raise my arm) has a physical effect within the brain, and in doing so initiates a causal chain ultimately leading to a particular bodily behaviour (the raising of my arm). Lowe rejects this proposal because he does not consider that there is a particular physical event that we can pick out as being the direct effect of the mental event, such that it would be correct to say that the mental event initiated a chain of causes that gave rise to the particular bodily behaviour. When the causal chain is traced back from a particular bodily behaviour that a mental event is assumed to have initiated, we find that the causal chains display a tree-like structure, the branches of which become entwined with other causal chains that result in other distinct bodily behaviours. Consequently, there is nowhere that one could place the mental cause within this entangled causal chain of physical events, such that it would be correct to say that this mental event initiates this particular bodily behaviour. (Lowe, therefore also rejects psychophysical reductionism as a possible form of physicalism as there is no particular physical event that a mental event could be identified with such that it would be correct to say that this event causes this particular

bodily behaviour. See Lowe (1996, pp. 66-67). Hence, according to Lowe although the 'causal tree *mediates* the causal relation between my desire or volition and the movement of my arm, it does not do so by virtue of my volition or desire initiating the 'growth' of the tree from the 'tips' down: for the tree has *no* tips, and certainly none that it can call exclusively its own' (1996, p. 65).

Mental events do not cause bodily behaviour by directly causing any physical event within the causal tree terminating in the bodily behaviour. Hence Lowe is willing to allow that the physical world is causally closed in the sense required by the argument from causal overdetermination — every physical event does have as a sufficient cause a physical event.⁶⁴ However, Lowe denies the conclusion of the argument from causal overdetermination because he denies the homogeneity of the causal relata. Mental and physical causation are causation by different kinds of entity. According to Lowe, a mental event causes a physical event not by initiating a sequence of physical events that result in this final physical event, but by bringing it about that a particular causal tree exists. A mental event is, in the circumstances, necessary and sufficient for the fact that this causal tree, that this particular kind of causal pattern of events terminating in the bodily behaviour, exists. This requires Lowe to admit that there is both event causation and fact causation, for according to this position mental events cause physical facts: A mental event causes the causal fact that 'certain physical events, P1, P2, ... Pn, are causes of another physical event, P' (2000, p. 16).

Although I consider Lowe's theory of psychophysical interactionism to have much to recommend itself, and to provide some crucial insights into interactive mentalism, I do not think that one can motivate a plausible distinction between event causation and fact causation, and hence allow Lowe's denial of the homogeneity of the causal relata. To be fair to Lowe, his aim in presenting this theory of interactive mentalism is to show that one possible way of responding to the argument from causal overdetermination is to deny the homogeneity of the causal relata. Indeed, the claim that there is fact causation is not one that I am sure he would want to admit within the context of his own ontological system.

⁶⁴ Lowe would in fact argue that even given the homogeneity of the causal relata, if the premise of causal closure within the argument from causal overdetermination is the premise that every physical event has a sufficient physical cause, this would still not allow us to identify mental and physical causes. This is because this causal closure claim is far too weak and is fully compatible with emergentism. The strength of the causal closure claim required within the argument from causal overdetermination will be discussed in detail in Part Three, § 1.2. Let it simply be said here, that Lowe's argument for interactive mentalism is compatible with a much stronger causal closure principle suitable for the argument from causal overdetermination.

However, even if for the purpose of the discussion one were not resistant to fact causation, I would suggest that there is no plausible ontological system which would enable one to combine fact causation with event causation in a way that would allow one to advance an interactive mentalism. In Part 3, §6.5, I argue that the central premises within Lowe's theory of psychophysical interactionism can be detached from Lowe's denial of the homogeneity of the causal relata. Here I only wish to consider the plausibility of Lowe's denial of the homogeneity of the causal relata.

It is not at all obvious how one should understand an 'event' and a 'fact' so that it is plausible to motivate a distinction between event causation and fact causation. One can, I think, lay down three plausible requirements for the denial of the homogeneity of the causal relata. Firstly, both events and facts must be suitable candidates for the causal relata. Secondly, event causation and fact causation, hence events and facts must significantly differ. Thirdly, the ontological considerations that motivate one's theory of events and one's theory of facts must be largely consistent with each other.

If there is both event causation and fact causation, then obviously events and facts must both be suitable candidates for the causal relata. This means that Lowe cannot be understanding a fact to be an abstract unity or to be an entity that is composed of purely abstract parts for these entities are more suited to the role of causal explanantia, and clearly Lowe wants the mental to have more than an explanatory role. If a mental event 'causes' a physical fact, then this must mean that mental causes have independent effects within the physical domain. Hence, for there to be fact causation, one must identify facts with states of affairs. So let us assume that facts are to be identified with Armstrong's states of affairs, and hence that a fact is a thin particular's instantiation of a property at a time. Furthermore, let us ignore the various problems with Armstrong's states of affairs. What is of concern here is whether, given this understanding of a 'fact' and hence 'fact causation', one can motivate a plausible but distinct understanding of event causation.

If one understands a fact in this way, how should one understand an 'event', so that it is plausible to motivate a distinction between event causation and fact causation? The obvious place to start is with Armstrong's understanding of an event. According to Armstrong, an event is a 'succession of states of affairs' (1997, p.206). An event is, in other words, a complex of states of affairs within a temporal zone. Now according to Armstrong, events are not entities that exist over and above states of affairs. Events reduce to complexes of states of affairs. But then if the only thing that distinguishes event

causation from fact causation is the fact that they relate different entities, then given Armstrong's understanding of an event surely this means that event causation is itself nothing over and above fact causation. It is true that causal statements that appeal to events will be less fine-grained, and thus less informative than those that appeal to facts, because the former will include complexes of facts. But I repeat that our concern is not with causal explanation but with causation. Given Armstrong's understanding of an event, event causation will be nothing over and above a complex of causal relations between facts.

If events are not successions of states of affairs, then what are they? To advance a distinction between event causation and fact causation, obviously one's theory of events and one's theory of facts must significantly differ. For this reason, one cannot combine a Kimean theory of events with an Armstrongian theory of states of affairs, for as discussed in Part One, § 5.2 it is most plausible that Kimean events are nothing other than Armstrongian states of affairs, and hence if one is understanding events in Kim's sense, event causation would be nothing other than fact causation, and mental causation nothing other than physical causation. In other words, if one identifies facts with Armstrong's states of affairs, and events with Kimean events, one is not really denying the homogeneity of the causal relata.

So the only option would seem to be to appeal to a theory of events which relates the ontological category of events to ontological categories which differ from the ontological categories to which Armstrong relates states of affairs. However, I cannot see how one could then make one's theory of events and one's theory of facts ontologically cohere with each other. To explain what I mean by suggesting that one's theory of events and one's theory of facts must ontologically cohere, consider the proposal that one combines an understanding of facts as Armstrong's states of affairs, whilst denying that events are related to the ontological category of property and hence maintaining a theory of events of the kind adopted by Davidson. The first problem with this is that, as is clear from our discussion, Davidson's and Armstrong's ontological systems are motivated by entirely different and incompatible considerations. Davidson's nominalism stems from the fact that his primary interest is in a semantics for natural language. Rejecting the truthmaker principle, Davidson does not need to appeal to properties to explain why events support certain descriptions rather than others because according to him there need not be anything about events that explains why events support certain descriptions rather than others (see Part One, § 2.5). On the other hand, Armstrong's motivation for advancing a

states of affairs ontology is his acceptance of the truthmaker principle. (See Part One, §5.3) And given the acceptance of the truth-maker principle it would seem that whatever one's understanding of an event, events must be related to the ontological category of property, for there must be some ontological ground which accounts for the fact that events support certain descriptions rather than others, and the ontological category of properties must be appealed to, to account for this. And, of course, one cannot consistently maintain the truthmaker principle in some ontological domains and yet not others. Rather, the truthmaker principle must serve as a guiding principle throughout one's ontological system.

Even if one ignores Armstrong's acceptance of the truthmaker principle, having accepted Armstrong's states of affairs and thus the ontological category of property, it is not plausible to suggest that events are the causal relata and yet that they are not related to the ontological category of property. All entities that causation relates must themselves be related to the ontological category of property, for as I have argued in Part One, § 2.5 and § 3.1, if there are properties they play the central role within all causation. The causal relations that an entity can enter into depend upon what that entity is like and hence which properties characterise it. Hence, if one admits the ontological category of properties, and some causation is a relation between events, I fail to see how events cannot be related to the ontological category of properties.

I hope that this shows that from an ontological point of view, at the very least it is not at all obvious how the interactive mentalist could plausibly respond to the argument from causal overdetermination by denying the homogeneity of the causal relata. Unless Lowe can provide an account of how one could do this, his solution to the problem of mental causation cannot be given any plausibility.

7

Summary

It is worthwhile summarising the main points of Part One of this thesis. After raising some general considerations about the causal relata in §1, in §2 I considered Davidson's anomalous monism and the commonly voiced criticism that it leads to a property epiphenomenalism. I distinguished between two kinds of property epiphenomenalism. The first understands the term 'property' in an ontological sense, while the second understands the term 'property' in a linguistic sense. I argued that even though Davidson's anomalous monism assumes a generalist theory of causation, it could not be accused of type epiphenomenalism (an epiphenomenalism of the first kind) because of Davidson's theory of events as the causal relata. Davidson does not relate the ontological category of events to the ontological category of properties. Nor is it plausible to accuse Davidson of predicate epiphenomenalism, because although it is consistent with Davidson's ontology, such accusations are ungrounded within Davidson's anomalous monism. However, I rejected Davidson's anomalous monism because I rejected his theory of the causal relata, and more generally his approach to ontology. One's ontological system and hence one's theory of the causal relata should not have semantic considerations at its base. If one's ontological system is not motivated by semantical considerations, but instead metaphysical enquiry, properties will inevitably play an essential role within one's ontological system and one's theory of causation. Hence, I arrived at my first positive claim — the causal relata are entities that have properties.

In §3, I argued that considerations of the qualitative specificity of causation and the transitivity of causation mean that causes must be property-instantiations. Mental causes cannot have epiphenomenal properties. This consideration does help to simplify the mental causation debate. There is no need to appeal to some sort of 'in virtue of' relationship — a relationship that is so open to misinterpretation. It also means that the argument from causal overdetermination is not an argument that can be raised at increasingly finer levels. It is not legitimate to ask about the causal efficacy of the 'mentalness' of a mental cause. If a mental cause is a cause, it is a cause full stop.

§ 4 and §5 were concerned with establishing what property-instantiations actually are, and how and to what degree this matters to the mental causation debate. I argued that it matters to the mental causation debate what the properties of causation are — that is, whether they are universals or particulars. This is not because I agree with Robb that an appeal to CPs as the properties of causation offers a plausible form of psychophysical reductionism that is able to avoid the problem of multiple realisability. On the contrary, I argued that one could not combine a trope monism with a type dualism in the way that Robb requires. Rather, what the properties of causation are, matters to the mental causation debate because what the criterion of existence and identity for properties is matters to the mental causation debate, but obviously to provide a criterion of existence and identity for properties one must first decide upon what properties are. Furthermore, I would suggest that the plausibility of the causal closure principle depends upon the theory of causation that is being assumed, which is in turn influenced by one's understanding of a property. Hence, how one approaches the mental causation debate greatly depends upon what one considers the properties of causation to be. I then went on to argue that what the properties of causation are, depends upon whether the causal relata are states of affairs or states of substances, which in turn depends upon what substances are. My conclusion was that given a plausible understanding of a substance, and the rejection of immanent realism, the properties of causation must be CPs.

Finally, in § 6 I considered whether it was plausible to deny the homogeneity of the causal relata in order to motivate a form of interactive mentalism in the way that Lowe does. I argued that the distinction that Lowe requires between event causation and fact causation could not be supported ontologically, and hence that it is implausible for the interactive mentalist to respond in this way to the problem of causal overdetermination.

Part Two

PROPERTY ANALYSIS

Property Analysis and the Mental Causation Debate

The causal relata are property-instantiations. Given that the causal relata are property-instantiations, I have argued that these are most plausibly interpreted as states of substances rather than states of affairs, and that this requires one to interpret the properties of causation as CPs rather than universals. A cause is a substance's instantiation of a CP.

The claim that the properties of causation are CPs rather than universals is crucial to one's analysis of the identity and existence conditions of causes and effects. This is because what it is for a cause (or effect) to exist, depends upon what it is for a property to exist, and what it is for two causes (or effects) to be identical with one another, depends upon what it is for one property to be identical with another property. But clearly what the *criterion of existence and identity for properties* is, depends upon what properties are.

What the criterion of existence and identity for properties is, therefore matters to the mental causation debate at the most general of levels, because what it is for a mental cause to exist depends upon what it is for a mental property to exist, and hence upon what it is for a property to exist. And equally, whether mental causes are identical with physical causes depends upon whether mental properties are identical with physical properties, which depends upon what it is for one property to be identical with another.

The further central issue within an analysis of properties, is the kind of relationships that properties can have with each other and other ontological categories. Of particular importance to the mental causation debate, and more specifically non-reductive physicalism, is the question of whether properties can be said to depend upon each other and, if so, in what sense. Hence, as well as considering what the criterion of existence and identity for CPs is, another central concern will be with CP dependence. The two issues are not entirely separate; a discussion of property dependence will, to some extent, be influenced by one's criterion of property identity.

Now to be a little more specific, I consider that a plausible analysis of properties reveals the true contenders within the mental causation debate. Importantly, I hope to show that it leads to the rejection of the popular idea that there is a non-reductive physicalist middle ground. There are many different kinds of non-reductive physicalism, which appeal to different kinds of dependence relationships between the mental and the physical in order to motivate their position. One can, I suggest, divide these various positions into two groups; those that appeal to a property layering within an object, and those that appeal to a property layering between objects. I purposefully tackle these two different forms of non-reductive physicalism separately, for they appeal to very different notions of property dependence. However, my general conclusion is that if the mental is related to the physical in the way that the non-reductive physicalist suggests, then given a plausible analysis of the existence and identity conditions for properties, far from allowing one to advance a property dualism, it should actually lead either to the identification of mental and physical properties (hence psychophysical reductionism) or the elimination of mental properties (hence eliminativism).

On the other hand, it is arguable that a plausible analysis of the existence and identity conditions for properties, severely weakens the argument from multiple realisability, hence increasing the plausibility of psychophysical reductionism. Hence, if one is going to be a physicalist, I would suggest that property analysis reveals that the true choice is between psychophysical reductionism and eliminativism. And given that one wants to maintain that there is mental causation, one's only option is to advance a psychophysical reductionism.

A familiar, but insufficiently respected point, that will be emphasised throughout this discussion, is the importance of distinguishing property-types from predicates. A plausible analysis of properties shows that not every meaningful predicate expresses a real property-type. Whilst few within the mental causation debate would admit to conflating predicates with properties, some of the central arguments and assumptions that are advanced, suggest otherwise. A failure to hold properties and predicates far enough apart, has led to the general conflation of the causal relata with the causal explanantia, and hence causation with causal explanation. In particular, the non-reductive physicalist appears to be guilty of confusing the idea that mental and physical predicates exist at different explanatory levels, with the idea that mental and physical properties exist at different ontological levels. Furthermore, once a proper understanding of the relationship between predicates and properties is gained, it should become clear that certain problems that have exercised those within the mental causation debate, for example, Kim's problem

of explanatory exclusion (Kim (1993e) and (1990)), are red herrings which should not distract those who are ontologically serious.

Before we can get onto a discussion of the ways in which an analysis of CPs affects the mental causation debate, we have an initial hurdle to get over. To raise the question of how one's criterion of CP identity affects the mental causation debate, is to assume that a criterion of identity can be offered for CPs. But a large number of philosophers have argued that the problem with CPs is precisely that they do not have determinate identity. If this was the case, then it would bring into question the central role that I have suggested that CPs play within causation. For this reason, our discussion begins with a defence of the claim that CPs have determinate identity and countability. What it is for one CP to be identical with another is a far from easy question to answer, and is an area that is relatively undeveloped within metaphysics; Certainly, Campbell (1991), the leading proponent of tropes, does not provide a fully satisfactory answer to this question. In the following, I argue that one CP is identical with another if it belongs to the same exact resemblance class and is instantiated by the same substance.

With this criterion of CP identity in mind, in §3 I move on to a discussion of supervenience and property dependence. This discussion begins with the observation that most formulations of supervenience assume a co-instantive supervenience. That is, a supervenience relationship in which the supervenient and subvenient property are both instantiated by the same object. Our discussion of this kind of supervenience moves from the now familiar consideration that co-instantive supervenience lacks the directionality of a dependence relationship, to the question of whether formulations of psychophysical co-instantive supervenience can plausibly be strengthened in order to capture this directionality. After reaching a negative conclusion, I then go on to consider the increasingly popular claim that although co-instantive supervenience is not a dependence relationship, a dependence relationship must be invoked in order to explain it.

In §4 our discussion returns to the analysis of CP existence and Identity. §4 and §5 are concerned with finding a plausible criterion of CP existence, and the question of when two CPs exactly resemble each other. Now Heil and Robb (forthcoming), reject any property layering within an object. Their argument is based upon the consideration that a semantical analysis of properties is incorrect. If Heil and Robb are correct, this will in fact lead to the rejection of all theories of non-reductive physicalism that consider there to be a relationship of co-instantive supervenience between mental and physical CPs. That is, it will lead to a rejection of all the positions discussed within §3. I agree that a semantical

analysis of properties is highly implausible (regardless of whether properties are universals of CPs) and that its rejection has a number of important effects within the mental causation debate. In particular, it leads to the rejection of Kim's Principle of Explanatory Exclusion, and arguably the rejection of the multiple realisability argument. More generally, it shows that from the fact that there are different levels of explanation one cannot conclude that there are different ontological levels of properties. However, contrary to Heil and Robb's suggestion, I argue that the rejection of a semantical analysis does not show that there are in fact no property levels within an object. To do this, one must advance a positive analysis of CP identity.

In §5, this takes us on to what I consider to be a more plausible analysis of CPs, namely a causal criterion. Given such a causal criterion, my conclusion is that all forms of non-reductive physicalism (not just those based within a co-instantive supervenience account) are implausible. However, this conclusion is not easily gained, and I suggest that the issue is rather more complicated than Kim's own discussion of this matter and his consequent rejection of non-reductive physicalism indicate (Kim 1992, 1993d, 1993g).

Our discussion of property analysis finishes with an examination of mereology. Up until this stage the discussion focuses on whether a single object can instantiate properties that are hierarchically ordered. But there is the distinct question of whether there is a hierarchical layering of properties between distinct objects. I consider the sense in which there could be a dependency relationship between mental properties and the properties of the physical objects that compose mental objects, and argue that this is in fact compatible with three conflicting positions within the mental causation debate. It is compatible with non-reductive physicalism, psychophysical reductionism, and full-blooded property emergentism. Because of our acceptance of a causal criterion of CPs, the first option must be rejected. And whether one advances a psychophysical reductionism or a full-blooded property emergentism, depends entirely upon whether one thinks that there is vertical causation, and hence whether one accepts the causal closure principle.

2

Do Characterising Particulars have Determinate Identity and Countability?

The causal relata are instantiations of a CP by a substance at a time, and in order to identify two causes they must involve the same CPs. Indeed, if two causes that occur at the same time, do involve the same CPs, then the causes must themselves be identical, as a single CP cannot be instantiated by more than one substance at a single time. (The same is not true of universals; the fact that two causes involve instantiations of the very same universal does not guarantee their identity, for the universal may be instantiated by different substances). Hence whether it is plausible that mental causes are identical with physical causes, depends upon whether it is plausible that a mental CP is the same CP as a physical CP.

But this is to assume that CPs have determinate identity. If there is no fact of the matter as to whether CP1 is the same CP as CP2, then although this should not necessarily lead one to banish CPs from one's ontology, it does call into question the central role that I have suggested that they play within causation, for surely there must be a fact of the matter regarding whether two causes or effects are identical. Worryingly, a large number of philosophers think that the problem with CPs is precisely that they lack determinate identity.¹

A further related question concerns whether CPs have determinate countability. There is a fact of the matter about how many chairs there are in this room, or how many pictures there on this wall. But is there a fact of the matter about how many green CPs this leaf instantiates or more generally how many CPs the set of exactly resembling green CPs contains? Once again, if CPs are to be the properties of causation, surely CPs must have determinate countability, because surely there must be a fact of the matter about how many distinct causes play a role in bringing about an effect.

¹ For example, see Lowe (1998, Ch. 3, pp. 79 -83).

Now there is no doubt that one can sensibly raise questions about the qualitative similarity and diversity of CPs. Hence, one can legitimately ask whether mental and physical CPs are exactly similar qualitatively, and thus belong to the same exact resemblance class. But from this one cannot conclude that CPs have determinate identity, as to suggest that all that is required for two CPs to be strictly identical is that they belong to the same exact resemblance class would be to collapse CPs into universals. The exact qualitative resemblance of two universals does entail their numerical identity.² If properties are universals, and the pain of this person and the pain of that person exactly resemble each other, then we are in the presence of a single property. However, with regard to CPs, exact qualitative resemblance should not be conflated with identity. Two CPs may not have a common nature, despite the fact that they have exactly resembling ones.

Because the exact qualitative resemblance of two universals entails their numerical identity, universals are easy to count. Counting CPs is difficult precisely because an exact resemblance class may contain more than one CP. Unless there is a fact of the matter about how many CPs there are within an exact resemblance class, CPs will lack determinate countability (and determinate identity). Although, it would be wrong to therefore conclude that a theory of universals has an advantage, for the question of how many instantiations of a single universal there are, parallels the question of how many distinct CPs an exact resemblance class contains.

One can tackle this problem by observing that what makes CPs from the same exact resemblance class numerically distinct is the fact that they are distinct particulars. Basing CPs within a two-category ontology, I have suggested that CPs are distinct particulars if they are instantiated by distinct substances. Hence, CPs from within a single exact resemblance class can be individuated from one another in virtue of the distinct substances that they are instantiated by. Therefore, for CP1 and CP2 to be strictly identical they must be instantiated by the same substance. Once again, this does not provide a sufficient condition for the identity of CPs; a single substance may instantiate CPs from within different resemblance classes, and hence distinct CPs. One should not identify the squareness of the rug and the redness of the rug, despite the fact that they are instantiated by the same substance. Nor obviously, should one identify a mental and a physical property merely on the basis that they are instantiated by the same person. Obviously, the solution is to combine the two claims. Two CPs are identical if and only if

² This is to assume what Armstrong refers to as the 'orthodox' analysis of resemblance, in which resemblance is understood in terms of having strictly identical properties. See Armstrong (1997, p.23).

they i) are instantiated by the same substance and ii) belong to the same exact resemblance class. Call this (ID).

Unfortunately, (ID) has counterexamples. Take the redness of this rug (which possesses exactly the same shade of red all over). From (ID) we should conclude that we are in the presence of a single CP, but this leads to problems. According to (ID), the redness of the LHS of the rug is not an individual CP, but only a part of one, for the redness of the LHS of the rug and the redness of the RHS of the rug both belong to the same exact resemblance class, and both are instantiated by the same thing — the rug, and hence should be identified. But the redness of the LHS of the rug can exist even if the redness of the RHS ceases to — if one were to dye the RHS blue, the redness of the LHS would not cease to exist. Indeed, given (ID), one must conclude that after the rug is dyed the LHS and the RHS of the rug instantiate distinct CPs, for the colour of the LHS and the RHS are no longer exactly resembling. Furthermore, if one were to wash the rug, thus removing the blue dye, one would once again be left with a single CP. Hence, (ID) entails that a single CP can become two numerically distinct CPs, and then can return to being a single CP. Worryingly, throughout this change, the original CP does not cease to exist. Although not logically impossible, this is metaphysically implausible. The cause of the single red CP becoming two distinct CPs was the application of the dye, which (unless we have a case of backwards causation) presumably occurred prior to the division of the single red CP. But then, if prior to the application of the dye, the redness of the LHS and the redness of the RHS were strictly identical, it follows that the dye would have to affect both CPs in exactly the same way, which it clearly need not.

The correct conclusion is that, even before it was dyed, the redness of the LHS of the rug and the redness of the RHS of the rug were numerically distinct, despite the fact that they belonged to the same exact resemblance class and were both instantiated by the rug. The redness of the rug is a complex of CPs, parts of which have persisted throughout the change, others of which have not.

Given these considerations one should conclude neither that (ID) is wholly false, nor that there is no fact of the matter regarding how many red CPs the rug possesses. Rather, the first part of (ID) — the idea of what it is to be a particular CP — needs refinement. The substance to which this counterexample appeals — the rug — is a composite substance, that is a complex substance having more basic parts. Hence, it is being assumed that a CP depends for its identity upon the composite substance that it is instantiated by. The rug's redness depends for its identity (and existence) upon the rug. The rug is what makes

it this redness, rather than some other redness, and this is because the particularity of a CP flows from the particularity of the substance that instantiates it. However, from the fact that substances make CPs the particulars they are, it is wrong to therefore assume that we should interpret substances to be composite. It is not the rug as a whole that makes the redness of the LHS of the rug the particular that it is. Rather, it is a part of the rug — the LHS. To this one might respond by introducing a further counterexample which parallels the one above: Is the redness of the LHS of the LHS of the rug the same CP as the RHS of the LHS of the rug? But fortunately, unless substances are infinitely divisible, this regress can be halted, for similar problems cannot be raised with regard to basic substances that have no parts. Hence, one may respond to this counterexample by advancing a criterion of particularity according to which a property is a particular if and only if it cannot be instantiated by more than one non-complex substance at a time. Two CPs are identical if and only if they (i) are instantiated by the same non-composite substance and (ii) belong to the same exact resemblance class. Call this (ID2).

But (ID2) would appear to entail that the only real CPs that there are, are those instantiated by non-composite substances. But then, what about the possibility that there are emergent properties that are instantiated by composite substances that are something over and above those instantiated by non-composite substances, mental properties arguably being one very important example? This kind of emergentism adopts the familiar layered model of reality, maintaining that some properties at the higher levels of the macro-micro hierarchy exist over and above properties at the lowest level. (Note that this understanding of emergentism is compatible with both non-reductive physicalism and full-blooded property-dualism, as it is neutral about whether emergent properties have independent causal powers). This kind of layered model of reality will be the subject of §6. The point that I wish to make here is simply that one's criterion of property identity should not rule out the very possibility of this kind of emergent property and yet (ID2) clearly does.

One can, I would suggest, accommodate the possibility of emergent properties by replacing talk of non-composite substances with talk of substances within a particular level on the macro-micro hierarchy. For example, call substances that are found at the third level on the macro-micro hierarchy of substances 'L3 substances', and properties that emerge at this level 'L3 properties'. An L3 property is not instantiated by substances at lower levels than L3, but may be instantiated by substances at higher levels than L3. Now let us say that an L3 property is a particular if it is not instantiated by more than one L3 substance at a time. It would then follow that two L3 CPs are identical if and only if they

are (i) instantiated by the same L3 substance and (ii) belong to the same exact resemblance class. Call this (ID3).

(ID3) is consistent with the possibility that there are emergent properties. However, it certainly does not entail that there are such properties — all properties may be 'Level One' properties. On the other hand, it does not encounter the kind of counterexamples that (ID) did. This is because two L3 CPs may belong to the same exact resemblance class and be instantiated by the same substance within a higher level of the macro-micro hierarchy and yet not be identical, as they may be instantiated by different L3 substances. Of the three criteria of identity for CPs, I would suggest that (ID3) is the most plausible.

Hence, let us assume that two exactly resembling CPs are strictly identical if and only if they emerge within the same level of the micro-macro hierarchy of substances (indeed in order for them to be exactly resembling they must emerge at the same level) and at that level are instantiated by the same substance. As CPs therefore have determinate identity and countability, the claim that the properties of causation are CPs is not threatened.

With regard to mental and physical CPs, we can therefore say that in order for them to be identical they must occur at the same level on the macro-micro hierarchy, and at that level be instantiated by the same substance. Furthermore, they must exactly resemble each other. A mental type is a set of exactly resembling CPs, that has members that have exactly resembling mental natures, which are instantiated by different substances. A physical type is a set of exactly resembling CPs that has members that have exactly resembling physical natures, which are instantiated by different substances.

Hence, as it should be, according to this criterion of CP identity, what it is that distinguishes a psychophysical reductionism from all the other positions within the mental causation debate that accept the existence of mental properties, is that psychophysical reductionism considers mental and physical CPs to exactly resemble each other. Some (but obviously not all) of the positions that accept a property dualism will agree with the psychophysical reductionist that mental CPs arise on the macro-micro hierarchy at the same level as some physical CPs, and at that level are instantiated by the same substance. However, they will deny that mental CPs exactly resemble physical CPs.

What therefore remains to be seen is when two CPs belong to the same exact resemblance class. Two CPs belong to the same exact resemblance class because of the

exact resemblance of their natures.³ Although one cannot analyse exact resemblance in terms of something else, it is quite legitimate to consider whether for two CPs to have exactly resembling natures they must exactly resemble each other in some further respect. Indeed, that one should raise such a question can be seen by the fact that it is not always easy to judge whether two properties do (exactly) resemble each other. For example, although we can all agree that scarlet things resemble crimson things more closely than they resemble blue things, are the types 'being triangular' and 'being trilateral' exactly resembling? And most certainly for the purpose of the mental causation debate, something more needs to be said about when one CP exactly resembles another. Do the types 'being mental' and 'being physical' exactly resemble? How should we decide upon such a matter? It is only once we have considered when CP1 belongs to the same exact resemblance class as CP2, that we can see what must be the case for a mental CP to be identical with a physical CP.

Now although, contrary to Robb (1997), the move from universals to CPs does not remove the problem of multiple realisation for psychophysical reductionism, the rejection of a certain implausible account of what it is for one CP to belong to the same exact resemblance class as another, does. The criterion that I am here referring to is a semantical one. The rejection of a semantical analysis of CPs will also bring into question the plausibility of the non-reductive physicalist's claim that there is a 'middle ground'. Of course, much of the physicalist's motivation for finding such a middle ground is removed, if the problem of multiple realisation is rejected. But a plausible analysis of properties also leads one to question the independent plausibility of non-reductive physicalism. Before completing my analysis of CPs, it is to this non-reductive physicalist middle ground which I turn. There are many different forms of non-reductive physicalism and it is arguably a particular family of them whose plausibility comes under threat once an adequate criterion of existence and identity for CPs has been grasped. To isolate this family of theories, it is first necessary to say something about non-reductive physicalism and its current status in the mental causation debate.

³ As discussed in Part 1, § 4.4.

3

Co-instantive Supervenience and Property Dependence

3.1 Supervenience and the Mental Causation Debate

Non-reductive physicalists maintain a property dualism, and like all physicalists are committed to the causal closure of the physical domain. However, we do not yet have physicalism; the epiphenomenalist, the parallelist, and indeed some forms of interactive dualism (e.g. Lowe's) are also committed to property dualism and are consistent with the causal closure of the physical domain. What the non-reductive physicalist needs to capture is the further idea that the physical determines the mental. Some attempt to do this by appealing to the idea that sets of mental properties 'supervene' upon sets of physical properties. A set of mental properties supervenes upon a set of physical properties if there can be no change in the mental properties without a change in the physical properties.

Nonreductive physicalists who appeal to the idea that mental properties supervene upon physical properties, hope to settle the ontological status of mental properties. Being a weaker relationship than identity, psychophysical supervenience is compatible with the claim that mental properties and physical properties are distinct, hence avoiding problems such as that of multiple realisability. Yet supervenience can allow for the primacy of the physical domain because it is thought to express an asymmetrical dependency relationship between the supervenient mental properties and the subvenient physical properties. Finally, with regard to the causal efficacy of mental properties, the non-reductive physicalist hopes to forge a route between epiphenomenalism and a 'full-blooded' interactive mentalism, despite denying the identity of mental and physical properties. As mental properties are wholly dependent upon physical properties, the causal powers of mental properties are wholly dependent upon or 'grounded in' those of physical properties.

There are two initial points that need to be made here. Firstly, the supervenience thesis is often said to be advanced in the context of a 'token identity theory'. At the token level, the mental and the physical are identical, but a relationship of supervenience exists between mental and physical types. Given a property-instantiation account of the causal relata, a supervenience thesis cannot be advanced in the context of an interesting token identity theory. This is because if the causal relata are property-instantiations, then if mental and physical types are distinct, mental and physical causes must also be distinct. This is equally true if the properties of causation are CPs rather than universals, because, as discussed in Part One § 4.4, contrary to Robb (1997), given a CP ontology, a token monism cannot be combined with a type dualism. If mental and physical types are distinct, mental and physical CPs must also be distinct.⁴ Therefore, the form of non-reductive physicalism that we are interested in here is one that maintains a substance monism together with a property dualism, and combines this with the claim that mental properties supervene upon physical properties. This is true, regardless of whether the properties of causation are universals or CPs.

Secondly, philosophers who support a theory of supervenience tend to assume that properties are universals. For example, consider the way in which the relationship of supervenience tends to be formulated. Weak psychophysical supervenience, for example, is normally considered to require that substances that are identical in regard to their physical properties within the same possible world are identical in regard to their mental properties.⁵ This clearly can only be the case if properties are universals, for different substances cannot instantiate identical CPs. Given a CP ontology, talk of universals must be replaced by talk of sets of exactly resembling CPs, and talk of identity between properties instantiated by distinct substances must be replaced by talk of exact resemblance between properties instantiated by distinct substances. Weak psychophysical supervenience must therefore be re-interpreted as the claim that substances that are exactly resembling in regard to their physical CPs within the same possible world, are exactly resembling in regard to their mental CPs. Presumably, a set of exactly resembling mental CPs supervenes upon a set of exactly resembling physical CPs, in virtue of the fact that each member within the first set supervenes upon a member

⁴ Note that, if, contrary to what I have argued, one could combine a CP monism with a type dualism, then whether or not mental types supervened upon physical types would not be of particular concern within the mental causation debate, because within a CP ontology it is most plausible that causation rests wholly upon entities at the token level.

⁵ For example, see Kim's definition of weak supervenience (1993i, p. 64) and his definition of strong supervenience (1993i, p. 65). Both assume that properties can be shared by different particulars.

within the second. To make matters more straightforward, in considering how to formulate the supervenience relationship in §3.2, I shall follow others in identifying properties with universals. However, when considering whether psychophysical supervenience can be plausibly strengthened in order to yield an asymmetry (§3.3), subtle differences between a supervenience relationship between CPs and a supervenience relationship between universals become relevant.

With good reason, the enthusiasm with which psychophysical supervenience was first met is now beginning to ebb. It is pointed out that the supervenience of one set of properties on another, merely presents a property correlation, and does not express a dependence relationship at all. Psychophysical supervenience is therefore equally compatible with various conflicting positions within the mental causation debate. Most worryingly, psychophysical supervenience is compatible with some forms of interactive mentalism. To be of any use to physicalism, psychophysical supervenience must be supplemented with some physically respectable explanation of why psychophysical supervenience obtains, although exactly what is contentious.⁶

One of the reasons why the inadequacies of psychophysical supervenience as the basis of a form of non-reductive physicalism has been overlooked for so long, is that supervenience's status as a dependence relationship between properties has been assumed without any proper investigation of the notion of property dependency. In order to explore the issue of property dependence one cannot simply lump all forms of non-reductive physicalism together. In §3.2, I argue that Kim, along with nearly all others attempting to formulate the supervenience relationship, assumes a co-instantive formulation of supervenience. That is, one in which the supervenient and the subvenient properties are instantiated by the same object.

In §3.3 and §3.4, I consider how a non-reductive physicalist who appeals to this kind of supervenience could attempt to establish a dependency relationship between mental and physical properties. In §3.3, I compare supervenient 'dependency' with standard ontological dependency. Standard formulations of co-instantive supervenience lack the directionality that ontological dependence relationships have. And after considering the ways in which ontological dependency is commonly formulated so that it captures an asymmetry, I argue that supervenience claims cannot be strengthened in a similar way.

⁶ On the need for explaining supervenience relationships see, for example, Kim (1993j), Heil (1998), and Horgan (1993)

Such adjustments are either inapplicable to the supervenience relationship or implausible in the case of psychophysical supervenience. In §3.4, I consider the argument that although co-instantive supervenience is not a dependence relationship, a dependence relationship between mental and physical properties must be invoked in order to explain it, and examine two of the physically respectable explanations that are typically appealed to.

3.2 Co-instantiation Theories of Supervenience

Discussions of supervenience have been complicated by the fact that there are so many nonequivalent formulations of the supervenience relationship, not all of which have the same non-reductive and determinative capacities. Numerous attempts to identify the different types of supervenience relationship and work out the logical relations among them have resulted in at least seven formulations that purportedly express a relationship that could ground the ontic priority claims that supervenience is supposed to entail. For the most part, the literature concerning psychophysical supervenience has concentrated on formulating these alternative conceptions of supervenience, questioning which, if any, is appropriate to the mind-body relation, and whether these are compatible with the non-reducibility of the mental to the physical.⁷ Here it is necessary to make only a few brief observations about the way in which supervenience is usually formulated, their point being to demonstrate that the various formulations are less disparate than is commonly considered. There is in fact a common thread which runs through many of them, which not only limits the kind of dependency relationship that they could involve, but which in the light of further property analysis should arguably lead us to abandon them all.

The difficulty for the non-reductionist in formulating supervenience is commonly thought to be that of finding a characterisation that is strong enough to meet the demands of asymmetric dependency without also leading to reducibility. The strength of a supervenience claim varies with the size of the subvenient domain and the strength of the necessity operator or the class of possible worlds that the particular formulation allows one to quantify over. Kim has distinguished between two types of supervenience, a weak and a strong version.⁸ In terms of possible worlds, weak supervenience only refers to substances that occupy the same possible worlds. It requires that substances that are

⁷ For example, see Kim (1993i), (1993k), McLaughlin (1995), Klagge (1995), and Post (1995).

⁸ See Kim (1993i, 1993j, 1993k).



identical in regard to their subvenient properties within the same possible world, are identical in regard to their supervenient properties, allowing discernibility of supervenient properties between substances that are indiscernible in regard to their subvenient properties across possible worlds. Strong supervenience requires, in addition, that if substances are subvenient-indiscernible across any two possible worlds, then they will be supervenient-indiscernible. As weak supervenience is silent about the dependence relationships that must obtain across possible worlds, it is arguably too weak to express a dependency relationship, because dependence needs to carry a modal force.⁹ Strong supervenience, on the other hand, appears to capture something that is closer to a dependency relationship as each supervenient property is necessitated by some subvenient property, although for this very reason there is controversy over whether strong supervenience is really a non-reductive relationship.¹⁰

A further distinction can be made between individualistic supervenience (according to which there are no two persons that are identical in all physical respects but distinct in some mental respect), regional supervenience (according to which there are no two regions that are identical in all physical respects but distinct in some mental respect) and global supervenience (according to which there are no two physically possible worlds that are identical in all physical respects but distinct in some mental respect).¹¹ What the relationship is between these various types of supervenience has been the subject of many debates. For example, Kim in 'Concepts of Supervenience' (1993i) maintains that global supervenience is equivalent to strong supervenience. In "'Strong' and 'Global' Supervenience Revisited" (1993k) he rejects this claim, arguing that global supervenience is weaker than strong. In 'Postscripts on Supervenience' (1993l) he suggests that global and strong supervenience are equivalent when restricted to intrinsic properties, equivalence only failing if extrinsic properties are present in the supervenient set and disallowed from the subvenient base.

This confusion can be removed once it is recognised that all of these theories are co-instantive theories of supervenience.¹² Despite the fact that one is referring to objects within a possible world and the other to objects across possible worlds, both weak and

⁹ See Kim (1993i, pp. 57 -64) for this argument.

¹⁰ On the question of whether strong supervenience is a reductive relationship, see Kim (1993i, §5). Also, see Savellos and Yalcin (1995, pp. 5-7).

¹¹ See Horgan (1993, §5) for a defence of regional supervenience, and Post (1995) for global supervenience.

¹² See Horgan (1995, §5) for further support of this claim.

strong supervenience draw the subvenient and supervenient properties from a single object. Global, regional, and individualistic supervenience disagree about what that object is. An individualistic theory of supervenience states that it is the same individual who is in pain and who is in c-fibre stimulation. With global supervenience the base has been widened to include all of the physical properties of a world, and the supervening properties have been widened to include all of the mental properties of the world, but this is still an example of a co-instantiation theory; the one major difference being that the object instantiating the supervenient and subvenient properties is a whole world rather than a single individual. Because the domain is far larger than that of an individualistic theory, this will generate a weaker theory of supervenience, albeit one that is still compatible with strong supervenience. Kim, along with nearly all others attempting to formulate the supervenience relationship, assumes a co-instantive formulation of supervenience. This common theme is crucial in assessing both their plausibility as providing a dependence relationship and their compatibility with non-reductive physicalism.

3.3 Property Dependence and Co-instantiation Theories of Supervenience

It is, or at least it was, commonly assumed that supervenience is a type of dependence relationship. The main problem for the non-reductive physicalist was thought to be that of articulating a type of supervenience that expressed a relationship weak enough to be compatible with the non-reducibility of mental properties to physical properties. Strong supervenience was generally considered to fit the bill. Therefore, once the physicalist had established that mental properties strongly supervene upon physical properties, it was assumed that he had also established that the mental properties depend upon physical properties. But can a co-instantive supervenience (whatever its strength) really constitute a dependence relationship?

One must first identify the kind of dependency supervenient 'dependency' is supposed to be. As psychophysical supervenience claims aim to capture a relationship between the ontological category of properties rather than propositions, it clearly cannot express a logical dependency. Given that the supervenience relationship is a relationship between properties, and therefore that the non-reductive physicalist is aiming to capture the idea that the physical is ontologically more basic than the mental, as Kim explains, 'what we want is *metaphysical* or *ontic* dependence' (1993j, p. 144). Perhaps, therefore, the

supervenience relationship aims to express some variety of dependency that is a close cousin to the more familiar ontological dependency.

Co-instantive supervenient 'dependency' is clearly not ontological dependency. Object s1 ontologically depends upon object s2, if what s1 is depends upon what s2 is. Most commonly, 'what something is' has been taken to concern the object's existence, thus leading to the claim that s1 ontologically depends upon s2 if and only if the existence of s1 depends upon the existence of s2. This has then been combined with a modal construal of dependence, to yield D1, where D1 is the claim that:

D1: s1 ontologically depends upon s2 if and only if it is necessary that s2 exists if s1 exists, where s1 is not identical with s2.¹³

The first important difference between the supervenience relationship and the relationship of ontological dependency is their relata. Ontological dependence is a relationship between objects, and supervenience is a relationship between properties. Are properties objects? Well, an 'object' is an entity possessing determinate identity conditions. Universals possess determinate identity conditions, and I have argued that the same is true of CPs, hence both qualify as objects. Even so, it is certainly not evident that properties ontologically depend upon each other. Certainly, if properties are CPs we can say that they ontologically depend upon the substances that instantiate them, that is, the existence of, for example, the rose's redness depends upon the existence of the rose. However, psychophysical supervenience is not a relationship between a property and the substance that instantiates it, but a relationship between properties. Now arguably some properties do ontologically depend upon each other. For example, the existence of compound properties depends upon the existence of the properties from which they are composed. However, this is a dependency relationship between properties belonging to two different objects — a part and its whole. This sort of property dependence is relevant to a different kind of supervenience — mereological supervenience — but the concern here is with co-instantive supervenience, and hence a relationship between properties instantiated by the same object. And it is certainly not immediately obvious that the existence of a property can ontologically depend upon the existence of another property instantiated by the same object as it.

¹³ This excludes self-dependence.

And there is a further fundamental difference. Standard ontological dependence claims are primarily concerned with what is required in order for one object to depend for its existence on another. They do not directly limit the amount of change that the dependent object can undergo, independently of the object upon which it ontologically depends. But what is crucial to supervenience is the relationship between variation in supervenient properties and variation in subvenient properties. Set of properties P1 supervenes upon set of properties P2, because P1 properties cannot vary without variation in P2 properties. Indeed, unlike standard ontological dependence, not all formulations of supervenience necessarily entail that in order for supervenient properties to exist their subvenient base must exist. Hence, according to the possible-worlds formulations of psychophysical supervenience (disallowing negative properties), physical properties do not have to exist in order for mental properties to exist. Two possible worlds would be indiscernible in respect to physical properties, if they both lacked physical properties and contained unchanging mental properties.¹⁴ Consequently, if co-instantive supervenience is a kind of dependency, it clearly is not ontological dependency.

But perhaps the supervenience relationship is some alternative dependence relationship, which belongs to the same family of dependence relationships as ontological dependence. This claim is implausible. When we consider ontological dependence relationships, we see that what is crucial to them all, is that, that which does the determining must, in some sense, be ontologically prior to that which gets determined by it. When objects s1 and s2 are non-identical, if s1 ontologically depends upon s2, then s2 cannot ontologically depend upon s1. But supervenience lacks this directionality. The strong supervenience of P1 on P2 neither implies nor precludes the strong supervenience of P2 on P1. It simply states a pattern of covariance between two properties.¹⁵ For example, imagine that two species x and y are symbiotic and that it is a law of nature that the number of x's and the number of y's are perfectly and positively correlated. Restricting the set of worlds to those in which this law obtains, the number of y's strongly supervenes on the number of x's. (And vice versa). However, it would be incorrect to suggest that these properties stand in a dependence relationship with one another, precisely because they do not stand in an asymmetrical relationship with one another. Supervenience cannot

¹⁴ It is for this reason that there are disputes about whether Kim's possible worlds formulation of supervenience and his modal formulation are equivalent. For details on this issue see Kim (1993k) and McLaughlin (1995)

¹⁵ See, for example, Savellos and Yalcin (1995, pp. 9-11), Kim (1993l), Kim (1999a, pp. 9-15) and Heil (1998) for this observation about supervenience.

purport to capture a dependence relation, because supervenience theses do not formulate asymmetric relationships.

It is because of this lack of directionality that it is increasingly recognised that psychophysical supervenience does not constitute an alternative theory of the mind-body relationship.¹⁶ It is in fact consistent with various conflicting positions, not all of which are non-reductive, and more worryingly not all of which are physicalist. For example, as identity is the strongest form of supervenience, psychophysical reductionism entails psychophysical supervenience. And, of far more concern for the physicalist, is that because of its non-asymmetry strong supervenience is compatible with 'full-blooded' mentalist positions.

Where should we go from here? Well, I would suggest that the non-reductionists' response should be to consider standard formulations of ontological dependence in the hope of finding out how supervenience claims might be strengthened so that they can also capture the required asymmetry. The remaining questions being whether such adjustments are applicable in the case of the supervenience relationship, and plausible in the particular case of psychophysical supervenience. And it is here that we hit an interesting problem, for it is questionable whether standard formulations of ontological dependence have had any more success in establishing asymmetry.¹⁷ Just as there is a problem in defining the supervenience relationship between properties so that it possesses an asymmetry, there is a problem in capturing the asymmetry between two objects within a definition of ontological dependency. Hence, returning to D1, interpreting 's1 ontologically depends upon s2' as necessarily s1 exists only if s2 exists, where s1 is not identical with s2, does not exclude the possibility that necessarily s2 exists only if s1 exists and therefore does not establish an asymmetry.

How do those concerned with formulating ontological dependency address this problem? Well, one way to modify the formulation of ontological dependency in order to capture the asymmetry is to define the ontological dependence of s1 on s2 in the following way:

D2: s1 ontologically depends upon s2 if and only if, necessarily s1 only exists because s2 exists, where s1 is not identical with s2.

¹⁶ For example, Kim (1993), Kim (1999a, pp. 9-15) and Heil (1998) argue that supervenience does not constitute an alternative position within the mental causation debate.

¹⁷ For example, see Lowe (1998) for a discussion of ontological dependency and the problem of capturing asymmetry within a definition of ontological dependency.

'Because' expresses an explanatory connection, its asymmetry arising from the fact that two separate facts cannot explain each other. But the problem with this suggestion is that we are seeking an ontological relationship of dependence, rather than an explanatory connection. For this reason, the solution is not appropriate to ontological dependence. Nor should it be of any interest to those wishing to formulate an ontological supervenience.

Alternatively, one might build an asymmetry into the ontological dependence relationship. Hence:

D3: s1 ontologically depends upon s2 if and only if, necessarily s1 exists only if s2 exists and it is not the case that necessarily s2 exists only if s1 exists, where s1 is not identical with s2.

Applied to supervenience, P1 depends upon P2, if P1 strongly supervenes upon P2, but P2 does not strongly supervene on P1. Given that supervenience is an expression of covariance, this can be analysed as P1 strongly covaries with P2, but P2 does not strongly covary with P1. However, in neither the case of ontological dependency nor of supervenient 'dependency', is this enough to establish an asymmetrical dependency, for even if the relationship went in one way only, then this would not automatically have established a dependence relationship. This is because the asymmetrical relationship between the two entities may be a consequence of a more fundamental relationship that they both have with a third entity. Furthermore, even if one could achieve asymmetry in this way, and hence the non-reductive physicalist advanced this strengthened formulation of supervenience in order to capture the dependence relationship between supervenient and subvenient properties, he would arguably have a pyrrhic victory if he tried to show that this strengthened form of supervenience relationship held true of the relation between mental and physical properties. According to Miller (1990), although empirical evidence may appear to support the claim that the mental strongly covaries with the physical but the physical does not strongly covary with the mental, this apparent asymmetry is merely an artefact of the contingent features of our physical and non-physical language.

There is, however, an alternative way of establishing an asymmetry within standard claims of ontological dependence. E. J. Lowe (1998, pp. 147-51) suggests that an asymmetry can be established if one appeals to identity dependence in one's understanding of ontological dependence. Hence:

D4: s1 ontologically depends upon s2, if and only if necessarily, the identity of s1 depends upon the identity of s2.

Lowe gives the example of the dependence of an assassination on the person assassinated. The identity of an assassination is partially fixed by the identity of the person assassinated. Consequently, the assassination cannot exist unless the person assassinated exists. Identity dependence is an asymmetrical relationship, because if the identity of s1 depends upon the identity of s2, then the identity of s2 cannot depend upon the identity of s1, if one is to give non-circular criteria of identity for s1 and s2.

Can this theory be applied to supervenience? In the first place, is this theory correct when applied to properties? Well, as our discussion of the identity of CPs has established, if properties are CPs, it is most plausible that CPs partially depend for their identity upon the identity of the substances that instantiate them. (The same is not true of universals. Even if universals are ways that substances are, although a universal will generically depend upon the substances that it is instantiated by, a universal does not depend for its identity upon the identity of these substances). But psychophysical supervenience is not a relationship between a property and the substance that instantiates it, but a relationship between two sets of properties instantiated by the same object. Does the identity of one property depend upon the identity of another? More specifically, does the identity of a mental CP depend upon the identity of the physical CP upon which it supervenes? It is not at all obvious that this is the case. However, one need not get too involved in this question, because Lowe's way of establishing asymmetry within ontological dependence is inapplicable to supervenience claims for the further following reason.

The fact that a supervenience claim, unlike standard ontological dependence claims, is not a claim about the way in which one property depends for its existence upon another, but instead is concerned with limiting the amount of change that one property can undergo independently of another, should lead one to question whether supervenience claims can be strengthened in this way. One would certainly have to modify this criterion of ontological dependence in order for it to be appropriate to supervenience. One would have to say something of the form: 'A set of supervenient properties superveniently depends upon a set of subvenient properties if a change in the identity of the set of supervenient properties depends upon a change in the identity of the set of subvenient properties.' Does this establish an asymmetry? That is, would it be viciously circular also to maintain that a change in the identity of the set of subvenient properties depends upon a change in the identity of a set of supervenient properties? It is certainly not clear to me

that this does carry the same sort of implication. I am therefore doubtful whether this way of establishing the asymmetry of ontological dependence is applicable to the supervenience relationship.

Consequently, after considering the ways in which ontological dependency might plausibly be formulated so that it captures the required asymmetry, it is not at all obvious that the non-reductive physicalist can plausibly modify his formulation of psychophysical supervenience in a similar way. Such adjustments seem to be either inapplicable to the supervenience relationship or implausible in the case of psychophysical supervenience. D2 is equally inapplicable to ontological dependency and to ontological supervenience because it expresses an explanatory connection. D3 also arguably fails to establish the required dependence relationship because the asymmetry may be a consequence of a third entity. Ignoring this problem, philosophers such as Miller would reject the claim that there is such a relationship between mental and physical properties. Finally, although D4 offers a plausible way of formulating ontological dependency, it is questionable whether the supervenience relationship can be strengthened in a similar way. In particular, this is because supervenience is primarily concerned with the relationship between variation in supervenient properties and variation in subvenient properties, while ontological dependence is primarily concerned with existence dependence.

3.4 Co-instantiation Theories of Supervenience with Supplementation

It is now increasingly urged that what is of interest to the mental causation debate is not that there is a relationship of supervenience between mental and physical properties, but *why* there is such a relationship.¹⁸ Papineau suggests that a lack of development in the question of *why* supervenience obtains 'has encouraged the view that the supervenience of the mental on the physical depends on some kind of basic intuition and that those who find physicalism unappealing are therefore free to reject it' (1995, p. 226). This is a curious comment, for, as we have seen, depending upon the particular anti-physicalism invoked, one may neither need, nor wish, to reject psychophysical supervenience. Indeed, a lack in the development of the question of *why* supervenience obtains, has encouraged the

¹⁸ For example, Kim suggests that supervenience "... reports a pattern of property covariation, suggesting the presence of an interesting dependency relation that might explain it." (1993, p.167). Also See Horgan (1993), esp. § 8 and Heil (1998).

physicalist to think that once an adequately formulated definition of supervenience has been provided, he need do no more work. On the contrary, given only the fact that the mental supervenes on the physical, the physicalist is not much further forward. To provide any support for physicalism, psychophysical supervenience must be explainable in a physically acceptable way, and even then, one will not have shown that the only, and indeed the best explanation, is a physicalist one.

Non-reductive physicalists have been quick to urge that although psychophysical co-instantive supervenience is not a dependence relationship, a dependence relationship between mental and physical properties can, and should be, invoked in order to explain it. Most commonly, they have appealed to the idea that mental and physical properties are related as disposition to categorical base, or as determinable to determinate to explain psychophysical co-instantive supervenience. According to the first position, mental properties are second order dispositional properties. That is, a mental property is the property of having some first-order property, which stands in certain causal relations to other physical, and mental properties. Hence, for example, the property of being in pain is the property of having some first order property, which is apt to be caused by bodily damage and apt to cause avoidance behavior. This is consistent with psychophysical co-instantive supervenience because first and second order properties are possessed by the same object. For example, the property of solubility instantiated by an aspirin is related to the aspirin's having a certain molecular structure. And equally, the second-order property of being in pain instantiated by a person is related to the person's having a first order property which realises it. It is also compatible with the multiple realisability of mental types, as in different physical systems, different first-order properties may occupy the causal role of a mental type such as pain. For this reason, this understanding of mental properties is most commonly thought to entail a non-reductive physicalism, for how can one identify a second order type with any of the first order types which multiply realise it?¹⁹

Yablo (1992) has done the most to defend the second position. Yablo assumes a non-reductive physicalism and adopts a strong theory of supervenience, according to which 'Necessarily, for every x and every mental property M of x, x has some physical property P such that necessarily all P's are M's' (1992, p. 254). Note that Yablo's formulation of supervenience is a co-instantive one as the subvenient and supervenient properties belong to the same object (x). Hence, it must be supplemented with some further physically respectable explanation of why psychophysical supervenience obtains.

¹⁹ The notable exception is Kim (1999a) who identifies second order properties with first order properties.

Yablo responds by appealing to the determinate / determinable distinction to characterise the distinction between supervening and subvening properties. Physical and mental properties are related as determinate to determinable. Clearly, this is consistent with a co-instantive theory of supervenience, because a determinate and its determinable are instantiated by the same object. For example, it is the same rose that is scarlet and red.

The determinate/ determinable distinction involves a notion of asymmetric dependence, because although determinates necessitate their determinables, determinables do not necessitate any particular determinate. Hence, for example, the fact that x is scarlet necessitates its being red, but something may be red without being scarlet. And for Yablo, it is asymmetrical necessitation that lies at the core of the determination relation (1992, p. 252). According to Yablo, if mental and physical properties stand in a relation of determination (or more precisely a relation of asymmetrical necessitation) then the fact that they stand in a relationship of supervenience would be explained. This is because Yablo considers that, given supervenience, if I have a mental property, then this means that I have a physical property, which guarantees that mental property. Likewise, determinates necessitate their determinables.

For it to be the case that mental properties are related to physical properties as determinables to determinates, it must be possible for a particular mental property to exist without the existence of the particular physical property that realises it. This particular aspect of the determinate/ determinable relation is explained by the fact that mental properties are multiply realisable (1992, p. 255).

Given multiple realisation and supervenience, 'it is a matter of necessity that something has a mental property iff it has a physical property by which that mental property is asymmetrically necessitated' (1992, p.256). Thus, consideration of psychophysical supervenience and multiple realisation allow one to conclude that mental and physical properties lie in a relation of asymmetrical necessitation and thus a relation of determination. This is explained if every mental property is a determinable, which has a range of physical properties as its determinates.

But there is a problem facing Yablo's account, and indeed all other non-reductive physicalist accounts that consider there to be a relationship of co-instantive supervenience between mental and physical properties. Consider Yablo's account. Mental and physical predicates may be related as determinable to determinate. But it is only if

determinable properties exist, and furthermore, it is only if they are distinct from determinate properties, that this would support non-reductive physicalism. If determinable properties do not exist, then if mental and physical predicates were related as determinable to determinate, then this would entail an eliminativism. If determinable properties were identical with determinate properties, then this would entail a psychophysical reductionism. Similar considerations apply if the mental and physical are related as a second-order property to a first-order property. Indeed, more generally, it is only if supervenient properties exist, and these are distinct from subvenient properties that this would support non-reductive physicalism. In order to address these questions we need to assess the relationship between predicates and properties, and to do this it is property analysis to which one must turn.

A Semantical Analysis of Characterising Particulars

4.1 A Semantical Analysis of Characterising Particulars and Property Layers

A single object can be attributively complex. For example, the red of this ball and the roundness of this ball are distinct CPs, and are both instantiated by the ball. But of the distinct CPs that a single object instantiates, are any hierarchically ordered? As we have seen, the non-reductive physicalist who supports a co-instantive theory of supervenience would clearly suggest that there are. All draw mental and physical properties from a single domain, and maintain that some kind of dependency relationship exists between them.

In 'Ontology and Mental Properties', Heil and Robb oppose any such property layering. Although they do not consider the view that 'properties are layered in the sense that properties had by an object are (often) distinct from properties had by parts of that object' to be objectionable, what they do object to is a particular type of property layering according to which '*one and the same* object can have many different layers of properties' (p. 10). Their central example of such a property layering involves determinate and determinable properties. With a red ball, for example, the 'properties being scarlet, being red, being coloured... are all taken to be distinct and possessed by the *same* object...' (p. 11). Contrary to this, there are no such property layers.

If correct, most obviously this would lead to the rejection of Yablo's non-reductive physicalist response to the problem of mental causation. If the mental and physical are related as determinable to determinate, mental and physical properties are not distinct. Hence, rather than non-reductive physicalism, if mental and physical properties do exhibit this kind of relationship, it in fact leads to a much stronger kind of physicalism. More generally, if there are no property layers within an object, this would lead to the rejection

of any non-reductive physicalist response based within a co-instantive theory of supervenience.

Heil and Robb's argument against the existence of such property layers is based upon the rejection of a semantical criterion of property existence and identity. To see why the rejection of a semantical criterion leads Heil and Robb to reject co-instantive supervenience, we must return to our analysis of CP existence and identity.

4.2 A Semantical Criterion of Characterising Particulars' Identity and Existence

A semantical criterion of property existence is implausible, as is the criterion of identity that follows from it. The strongest kind entails that properties are nothing more than the meanings of predicates. In terms of universals:

E1: Universal U1 exists if and only if predicate *p1* exists.

How should E1 be understood given an ontology of CPs? It might be assumed that E1 does not sit well with an ontology of CPs, for predication is a process in which the same thing is being attributed to different substances. Indeed, one may even advance a linguistic version of the problem of universals: How can a single general term apply to an indefinite number of substances? Universals can be postulated to serve as the meanings of general predicates. Hence, when the predicate 'redness' applies to more than one substance at the same time, both substances will have the universal property of redness. But what of CPs? One should respond by appealing to sets of exactly resembling tropes. E1 should then be interpreted as the claim that:

E1*: The set of exactly resembling CPs that is Type-CP1 exists if and only if predicate *p1* exists.

As sets depend for their existence on the existence of their members, given E1* it follows that if the predicate *p1* exists, the set of exactly resembling CPs that is of Type-CP1 is not an empty one. At least one CP must exist that is of Type-CP1. The following criticisms of

a semantical analysis of properties shall be expressed in terms of CPs, but they also apply if properties are in fact universals.²⁰

E1* is implausible, not least because it conflicts with immanent realism, according to which:

E2: CP1 exists if and only if CP1 is instantiated.

Contrary to E1*, E2 entails that there are predicates for which there can be no sets of exactly resembling CPs. For example, the predicate 'being both square and round at t' cannot pick out a set of exactly resembling CPs, as it is logically impossible for the CPs that are its members to be instantiated. Furthermore, as Lowe (1999b) argues, the very idea that the predicate 'x does not instantiate itself' could pick out a property-type leads to contradiction. Hence, contrary to E1*, it is clearly not the case that for every distinct predicate there exists a corresponding set of exactly resembling CPs. Given that not every meaningful predicate expresses an existing property-type, when does a meaningful predicate express a property-type?

E2 cannot provide a criterion of CP existence. Certainly, E2 allows one to form conclusions regarding the existence of a particular CP; the red of this rose exists because the rose is red, 'being square and round at time t' does not refer to a set of exactly resembling CPs because no substance could instantiate a CP from this set, and mental CPs exist if and only if there are mental substances. However, as the right hand side of E2 quantifies over CPs it cannot provide a criterion of existence for CPs in general. But given E2, our search for a criterion of CP existence does become more focused. If an informative criterion can be given for deciding which CPs substances do instantiate, then given E2, this just is to give a criterion of CP existence.

One way of doing this is to appeal to an alternative semantical criterion. Namely:

E3: s1 instantiates CP1, if and only if s1 falls under the predicate *p1*.²¹

Hence, the CP that is the rose's redness exists if and only if the predicate 'red' is predicable of the rose. But although E3 is compatible with E2, for it only tells us what must

²⁰ See, for example, Armstrong (1997, pp. 25–28) who argues against a semantical analysis of universals.

²¹ This is what Armstrong refers to as predicate nominalism, which he rejects. See Armstrong (1989, pp. 10–11).

be the case for a substance to instantiate a CP and is silent about the ontological significance of those predicates that a substance cannot fall under, no one who treats the category of properties with ontological seriousness can accept E3, for a property's existence does not depend upon it having a semantic role. E3 has got it the wrong way around. Predicates do not determine the truth-value of instantiation claims, rather it is properties that make statements about predicates true, for it is properties which give predicates their meaning. And they can give predicates their meaning precisely because they do not ontologically depend upon them. Hence, just as there are predicates for which there can be no sets of exactly resembling CPs, it is possible that there are sets of exactly resembling CPs for which there are no predicates, not least because it is metaphysically possible and empirically probable that there are CPs which are as yet undiscovered.

Although the fact that sets of exactly resembling CPs can exist independently of predicates means that it is false that s_1 instantiates CP1 if and only if s_1 falls under the predicate p_1 , it is consistent with E4, where E4 is the claim that:

E4: For each distinct predicate p_1 that a substance falls under, there exists a single set of exactly resembling CPs – Type-CP1, and if s_1 falls under p_1 it is because s_1 instantiates CP1 (a CP from the set Type-CP1).²²

Hence, if s_1 falls under the predicate 'red', then there must be a set of exactly resembling CPs that are of the type red, and s_1 must instantiate a CP from this set. As with E3, E4 is compatible with E2, for its only concern is with those predicates that a substance falls under, but unlike E3, E4 does not entail that it is necessary for p_1 to exist in order for the set of exactly resembling CPs that is Type-CP1 to exist. What E4 does provide is a sufficient condition for CP existence. If a substance falls under a predicate, there exists a corresponding set of exactly resembling CPs and the substance instantiates a CP from this set.

Furthermore, E4 entails that if s_1 and s_2 both fall under the predicate p_1 , both instantiate CPs from the same set of exactly resembling CPs. *This apple and this rose both fall under the predicate 'red'*, so according to E4, the red of this apple and the red of this rose exist, and these CPs exactly resemble each other. Hence, in response to the question posed at the end of §2, namely when do two CPs exactly resemble each other, one can respond

²² See Armstrong (1972, p.164) for this kind of criterion of property existence. Note that Armstrong raises it, in order to reject it.

that it is fixed by the meaning of predicates. Synonymous predicates designate the same set of exactly resembling CPs, and non-synonymous predicates designate distinct sets of exactly resembling CPs. Hence, let us say that *s1* falls under the predicate *p1* in virtue of the fact that it instantiates CP1, and *s2* falls under the predicate *p2* in virtue of the fact that it instantiates CP2. If *s1* is the same substance as *s2*, and *p1* is synonymous with *p2*, then CP1 is identical with CP2.

In support of E4 it may be argued that in order for the predicate 'red' to apply to the apple, or at least for it to meaningfully apply, the predicate 'red' must designate a single set of exactly resembling CPs which corresponds to it. But taking an argument advanced by Armstrong (1989, ch.5,§3) about universals and transposing it onto CPs, consideration of predicates such as 'game' suggest that this is in fact false. Unless one admits wildly disjunctive properties, there is no set of exactly resembling CPs that all and only games have in common, in virtue of which they are games. But although 'game' does not denote a single set of exactly resembling CPs, this does not mean that it is incorrect to describe anything as a 'game', or that this predicate lacks meaning. The predicate 'game' applies to different substances in virtue of inexact similarities between them. Hence, let us say that A, B, C, D, E, and F are six different sets of exactly resembling CPs. Game W instantiates CPs from ABC, game X instantiates CPs from BCD, game Y instantiates CPs from CDE, and game Z instantiates CPs from DEF. Although there is not a single set of exactly resembling CPs that all of them instantiate CPs from, it may be correct to describe each as a 'game' in virtue of the different partial resemblances existing between them.

Hence, although if *p1* holds of *s1*, it is because of a CP or a complex of CPs that *s1* instantiates, contrary to E4, for '*s1* is *p1*' to be meaningful it is not necessary for the set of exactly resembling CPs that are Type-CP1 to exist, because it is not necessary for *s1* to instantiate CP1. Nor need there be any other set of exactly resembling CPs, members of which all things that are *p1* must instantiate. Rather, the predicate *p1* may apply to different substances in virtue of different types.

These considerations lead not only to the rejection of E4, but the criterion of identity that goes with it. If a single predicate does not correspond to a single set of exactly resembling CPs, nor will a predicate that is synonymous with it. With the rejection of E4 what of the claim that the non-synonymy of predicates leads to property distinctness? Respecting the fact that predicates need not designate a single set of exactly resembling CPs, one could modify the claim that if two predicates are non-synonymous then they designate two distinct sets of exactly resembling CPs, to the claim that if two predicates are non-

synonymous then they designate two distinct disjunctive sets of exactly resembling CPs. However, this fails to appreciate the more general point that a difference in meaning need not signify an ontological difference, for the identity of a CP does not depend upon its semantic role. Indeed, the fact that the non-synonymy of predicates does not entail distinctness of properties can be seen by the numerous counterexamples in which one (disjunctive) set of exactly resembling CPs has corresponding to it non-synonymous predicates. 'Is red' and 'is the colour of this apple' may involve exactly the same sets of exactly resembling CPs, as may 'being triangular' and 'being trilateral', or 'temperature' and 'mean molecular kinetic energy.' Attempts to explain these counterexamples away (e.g. Tye (1982)) fail to respect the true relationship between predicates and properties.

4.3 Applications within the Mental Causation Debate

A semantical criterion of CP analysis often appears to be at work within the mental causation debate. Philosophers of mind have rarely been careful to distinguish properties from predicates. Indeed, it is occasionally suggested that for the purpose of the mental causation debate it does not really matter whether one talks about predicates or properties. Hence, for example, in Heil's earlier work on the mental causation debate, he considers such a matter not to make a difference to the mental causation debate, although admitting to having 'a slight preference for thinking of these matters in terms of predicates', largely because he lacks 'a clear conception of what exactly properties are, how they are known, and how they are individuated.' (1992, p. 19). The reason why Heil is able to take such a relaxed attitude to the property-predicate distinction is because he considers that a 'property is exemplified whenever a predicate associated with it is satisfied.' (p.19) — A criterion of property existence along the lines of E4.²³

i) Explanatory Levels

The acceptance of E4 (along with any of the stronger semantical criteria of property existence) does have serious consequences within the mental causation debate. In the first place, it narrows down the possible physicalist stances that one may plausibly take in response to the problem of causal overdetermination. Both eliminativism and psychophysical reductionism appear highly implausible if one is assuming a semantical

²³ Clearly, Heil has since changed his mind about property analysis, and the importance of it within the philosophy of mind.

analysis of properties. Consequently, given a semantical analysis of properties, it seems that the only physicalist option is to accept a non-reductive physicalism.

One way to respond to the problem of causal overdetermination is to accept eliminativism, hence denying the premise of psychophysical causation. But given E4, the fact that 'being in pain' is a meaningful predicate that can be correctly ascribed to *s1*, means that this set of exactly resembling mental CPs must exist, and *s1* must instantiate a CP from this set. In denying the existence of mental properties, given a semantical analysis of properties, the eliminativist is forced to maintain that it is incorrect or meaningless to ever describe a person as in pain.

We can see that eliminativism does not lead to such a counter-intuitive conclusion, if a semantical analysis of CPs is rejected. From the fact that a substance falls under a mental predicate it does not follow that it instantiates a mental CP — a thing which can make a causal difference in the physical domain. Although it is true that a substance falls under a mental predicate because of the CPs that it instantiates, these need not be mental CPs. This is not to suggest that eliminativism is a position that one should adopt. The eliminativist's claim that there are no mental CPs and hence no mental causation remains implausible. My point is simply that the further objection that, given eliminativism, mental predicates cannot be meaningfully ascribed to substances is false, given the rejection of a semantical criterion of CP analysis.

It is also arguable that a semantical analysis of properties has played a role in the demise of psychophysical reductionism. Psychophysical reductionism is unpopular amongst physicalists because of the argument from multiple realisability. But with Heil and Martin (1999), I would argue that the argument from multiple realisability loses its credibility once one has moved away from a semantical analysis of properties. In terms of CPs, the multiple realisability argument rests upon the consideration that two substances instantiating CPs from entirely distinct sets of exactly resembling physical CPs, both fall under the mental predicate 'x is in pain' and hence both instantiate the mental type pain. It therefore follows that the mental type is multiply realised by the physical types, and hence cannot be identified with any of them.

To generate the multiple realisability argument, the assumption therefore appears to be that there must be a set of exactly resembling pain CPs corresponding to the predicate 'x is in pain' and that every substance that falls under the predicate 'x is in pain', must fall under it because it instantiates a CP from this set of exactly resembling CPs. Hence, from

the fact that a human and an octopus both fall under the predicate 'x is in pain', we can conclude that they share the mental type of being in pain designated by the predicate 'x is in pain.'

Indeed, the links between a semantical analysis of properties and the argument from multiple realisability is evident when one considers the way in which the argument is usually explained. To quote Putnam 'if we can find even one psychological predicate which can clearly be applied to both a mammal and an octopus (say "hungry"), but whose physical-chemical "correlate" is different in the two cases, the brain state theory has collapsed' (1980, p. 228). Similarly, Bickle defines the multiple realisability as the claim that: 'Mind cannot reduce to brain, nor psychology to neuroscience, because mental kinds are multiply realizable at all physical levels of description' (1996, p. 59). And Horgan, considers that mental properties are 'reducible to physico-chemical properties, and ultimately to microphysical properties. Reducibility, as usually understood, involves...the definitional equivalence... of property-expressing predicates....' (1982, p. 29). (*My emphasis*).

Given the rejection of a semantical analysis of properties, one can respond to the multiple realisability argument by pointing out that from the fact that various substances fall under the predicate 'x is in pain', it does not follow that they all instantiate exactly similar mental CPs, and hence it does not follow that mental types are multiply realised. The type 'pain' may function rather like the type 'game' in that 'pain' does not denote a single set of exactly resembling CPs. Although each subject that satisfies the predicate 'x is in pain' instantiates CPs in virtue of which it is correct to describe it as in pain, the CPs need not be the same in every case; there need not be a set of exactly resembling CPs that each of these substances instantiates CPs from. Therefore, what property it is in virtue of which a being is in pain could vary widely across species, or indeed individuals. It is, however, correct to describe each as 'in pain' in virtue of the different partial resemblances existing between them. Of course, mental predicates are multiply realised, because different substances fall under 'x is in pain' in virtue of their instantiating different properties, but, clearly, given our rejection of E4, this should not serve to hinder the identification of mental and physical CPs.

It is therefore open to the psychophysical reductionist to argue that the predicate 'being in pain' applies to different substances in virtue of inexact physical similarities between them. Certainly, as the non-synonymy of predicates does not entail the distinctness of properties, from the fact that mental and physical predicates are non-synonymous one

cannot infer that mental and physical types are distinct. Rather, the CPs in virtue of which it is correct to describe a substance as mental, may be the same CPs in virtue of which it is correct to describe it as physical. In other words, a single CP may have both mental and physical predicates corresponding to it. Those who support the multiple *realisability* argument could respond that some mental type which does denote a single set of exactly resembling CPs is multiply realised by physical types, but the onus is on them to show that there is any such case. At the very least, given the rejection of a semantical analysis of properties, the argument from multiple realisability loses much of its persuasive force.

Where does the rejection of a semantical criterion leave non-reductive physicalism? Certainly, the recognition that a semantical analysis of properties is false, removes some of the motivation behind non-reductive physicalism, because one of the main barriers in the way of accepting a psychophysical reductionism — that of multiple realisability — has *arguably been removed*. But does the rejection of a semantical criterion affect the independent plausibility of non-reductive physicalism itself? If Heil and Robb are correct, then the rejection of a semantical criterion leads to the direct rejection of all those forms of non-reductive physicalism that assume a co-instantive supervenience.

To introduce Heil and Robb's objection to this kind of non-reductive physicalism, let us compare the kind of supervenience that these non-reductive physicalists appeal to, with the kind of supervenience that Davidson appeals to within his account of the mental. These types of supervenience differ with regard to the kind of entity that they relate. Davidson (1980d, 1993) considers supervenience to be a relationship between predicates. This results in an ascriptive supervenience; there are various levels of explanation and the relationship between these levels is one of supervenience, subvenient predicates having greater explanatory power or generality than supervenient predicates. Ontological supervenience, on the other hand, is a relationship between properties and entails that there are various ontological levels. In this form, if the mental supervenes upon the physical, this means that there are mental and physical properties, and these are found at different ontological levels, physical properties occupying a lower level than mental properties.

Now I would suggest that much of the critical discussion about Davidson's theory of supervenience can be ignored, for his critics mistakenly assume that he is referring to an

ontological supervenience.²⁴ This relates to my original point that Davidson's critics have misunderstood his ontological system. When Davidson states that supervenience is a relationship between properties, his critics mistakenly interpret his use of the term 'property' in an ontological rather than a linguistic sense. However, if one assumes a semantical criterion of property existence and identity, then the distinction between ascriptive and ontological supervenience is lost anyway. For example, if E4 was correct, then from the fact that there are various levels of predicates, it would be natural to infer that there are various levels of properties to which these predicates correspond; for each distinct referring predicate there is a property, and as the predicates are hierarchically ordered one can infer that the properties to which they correspond are similarly ordered. Hence, given E4, the existence of explanatory levels leads to ontological levels. Supervenience is appealed to, to explain the relationship between these levels.²⁵

A failure to distinguish between ontological and ascriptive supervenience, and the more general failure, that Heil and Robb discuss, of unconscious switching from talk of explanatory levels to talk of ontological levels, is indicative of the fact that a semantical criterion of property existence and identity is implicitly being assumed within the mental causation debate. Without a semantical criterion of CP analysis, the fact that there are different levels of explanation does not imply the existence of distinct ontological levels. Due to their non-synonymy mental and physical predicates arise at different levels of explanation, but this does not entail that mental and physical properties arise at different ontological levels.

On this basis, Heil and Robb go on to reject any property layering within an object. Thus returning to Yablo's (1992) account, if the mental and the physical are related as determinable to determinate, then what this implies for the mental causation debate depends, not only upon how the relationship between determinable and determinate is

²⁴ For example, Kim's criticisms of Davidson are only applicable within the context of ontological supervenience. (1993h, pp. 22-23). Equally, McLaughlin argues that Davidson's supervenience is logically equivalent to Kim's weak supervenience, despite the fact that Davidson's is a relationship between predicates and Kim's is a relationship between properties (1993, p. 36).

²⁵ Interestingly, in his discussions of supervenience, Kim (1993g) understands the term 'property' in a very liberal way, far more liberal anyway than that which should be allowed by Alexander's Dictum — a theory of property existence which is at the core of Kim's rejection of non-reductive physicalism. Within *Concepts of Supervenience*, Kim assumes that sets of properties are closed under Boolean property-forming operations, complementation, and disjunction (also possibly infinite disjunction and conjunction) (1993i, p. 58). From this one might conclude that with regard to the supervenience relationship, by the term 'property' Kim means 'predicate', entailing that his concern is really with ascriptive supervenience.

understood, but upon what the relationship is between. Yablo clearly considers the relationship between determinables and determinates to be a relationship between properties, where determinable properties are distinct from determinate properties. But in doing so, Heil and Robb would argue that he is conflating properties with predicates. The predicate 'red' is a determinate of the predicate 'colour' and a determinable of the predicate 'scarlet'. If a semantical criterion was assumed, one could infer from this that both determinate and determinable properties existed and furthermore that due to their non-synonymy they must be distinct. Hence, the red of this ball is a CP and is related to a distinct CP — the colour of this ball, as determinate to determinable, and as determinable to determinate to a further CP — the scarlet of this ball. But Heil and Robb appeal to the implausibility of a semantical criterion of property identity and existence, in order to reject any such property layering. The colour of this ball, the redness of this ball and the scarlet of this ball are not three distinct properties. Hence, if Yablo is correct to argue that the mental and physical are related as determinable to determinate, this in fact runs counter to any account which acknowledges the existence of mental properties whilst denying their identity with physical ones, and thus it runs counter to non-reductive physicalism. Generalising this objection to all property layering within an object, a non-reductive physicalism that assumes a co-instantive supervenience collapses non-reductive physicalism into psychophysical reductionism.

I consider the rejection of a semantical analysis to remove much of the motivation for adopting non-reductive physicalism. However, contrary to Heil and Robb's indications, I do not think that, given the rejection of a semantical analysis of properties, one can automatically reject all property layering within an object. Heil and Robb's argument against the existence of ontological layers within an object needs further consideration. They move from the rejection of a semantical analysis of CPs to the rejection of all property layering within an object. This is an unjustified jump. The rejection of a semantical account shows that, for example, from the fact that determinate predicates and determinable predicates are non-synonymous predicates which a single object falls under, one cannot thereby infer that determinate and determinable properties are distinct, and hence that there are property layers within a single object. More generally, from the fact that there are different levels of predicates that a single object falls under, one cannot conclude that there are different ontological levels of properties within an object. But contrary to Heil and Robb's indications, it does not show that there are in fact no property layers within an object. To do this, one must advance a positive analysis of CP identity (which Heil and Robb do not do). Before I go on to consider whether a positive account of property analysis does in fact lead to the rejection of those forms of non-reductive

physicalism that appeal to psychophysical co-instantive supervenience, I want to consider one final way in which the implicit acceptance of a semantical criterion has served to confuse the mental causation debate.

ii) Kim's Principle of Explanatory Exclusion

The implicit acceptance of a semantical criterion has served to muddy the water with regard to exactly what causal problems one must face within the mental causation debate. Here I have in mind Kim's principle of explanatory exclusion. This states that:

EE: There can be no more than a single complete and independent explanation of any one event.²⁶

If this is combined with the 'Principle of Causal Explanatory Closure', according to which every physical occurrence is entirely explicable by reference to comparable physical occurrences, then unless mental explanations just are physical explanations (or the former are in some way dependent upon the latter), there is no room for mental explanations of physical events. For example, Fred raises his arm in order to vote. A mental causal explanation would cite his intention as a cause of his arm raising. A neuro-physiological causal explanation of his arm raising would say nothing about his intention, but would connect his arm raising with the firing of certain neurones. Given Kim's Principle of Explanatory Exclusion, either the two explanations are identical or they are inter-dependent or when considered individually, contrary to the Principle of Causal Explanatory Closure, they are incomplete.

Kim's argument is structurally similar to the argument from causal overdetermination. Given the denial of systematic causal overdetermination then:

CO: There can be no more than a single complete and independent cause of any one event.

The Causal Closure Principle states that every physical effect has a sufficient physical cause. And so together CO and Closure entail that, unless mental causes are physical

²⁶ For his principle of explanatory exclusion see (Kim 1993e) and Kim (1990).

causes (or the former are in some way dependent upon the latter), there is no room for psychophysical causation.²⁷

Does the Problem of Explanatory Exclusion impose a constraint upon theories of mental causation that is additional to that imposed by the argument from Causal Overdetermination? I do not consider that the Problem of Explanatory Exclusion can be plausibly interpreted in such a way that it does. In fact, either Kim's Problem of Explanatory Exclusion just is the Problem of Causal Overdetermination or it should be rejected. More specifically, unless EE and CO are equivalent, and hence by a 'causal explanation' Kim really means what is normally referred to as a 'cause', EE should be rejected.

According to Kim, 'whatever model of explanation you accept, unless you take a wholly fictionalist view of explanation, the principle of explanatory exclusion is plausible.' (1993e, p250). However, in defending EE, Kim assumes an explanatory realism. According to this model of explanation, 'to have a causal explanation of an event requires that the event specified as its cause be, in reality, a cause of that event.' Where e1 and e2 are events 'a causal explanation of e1 in terms of e2 is a "correct explanation" only if e2 is in reality a cause of e1' (1993e, p. 256). There must be, in other words, some objective basis — a causal relation — that grounds a correct causal explanation.

Contrary to Kim, I consider that EE is implausible, regardless of the model of explanation that one is basing it upon. It should be rejected even if one assumes explanatory realism. To suggest otherwise is ultimately to assume a semantical criterion of property identity and existence.

The causal relata are ontological entities. Causal explanation, on the other hand, is a relationship between linguistic entities, most plausibly propositions. Hence the causal explanantia have predicates, not properties, as constituents. Given explanatory realism, the correctness of the causal explanation: 'It is because Fred believed that candidate X was the best, that he voted for candidate X' depends upon whether the properties in virtue of which it is correct to describe Fred as having this belief are the cause of his voting for candidate X. More generally the causal explanation: 'It is the event that is the instantiation

²⁷ Philosophers of mind often appeal to something like Kim's Principle of Explanatory Exclusion. For example, Jackson (1996) appeals to the Principle of Causal Explanatory Closure (p. 378) and seems to conflate it with the causal closure principle. This conflation is explained by the fact that he considers properties to figure in both causation and causal explanation. (In particular, see fn. 19.)

of $p1$ by $s1$, that caused the event that is $s2$ instantiating $p2'$ is correct only if the predicates $p1$ and $p2$ truly hold of $s1$ and $s2$ respectively, which wholly depends upon the properties that the substances instantiate. For the causal explanation to be true, it must be the properties of the two substances in virtue of which $p1$ and $p2$ hold that make the causal difference.

Given these considerations one can see that explanatory realism provides a necessary condition for the individuation of explanations. Assuming explanatory realism, if two explanations refer to distinct events, then the explanations must themselves be distinct. Hence, explanatory realism together with CO, entail that if two causal explanations cite distinct and independent events as complete causes of the same effect, then one of the causal explanations must be false. So if explanation E1 cites event $e1$ as a complete cause of $e3$, and explanation E2 cites event $e2$ as a complete cause of $e3$, and $e1$ and $e2$ are distinct and independent, then E1 and E2 cannot both be correct causal explanations of the effect $e3$.

But the combination of explanatory realism and CO does not entail that there cannot be more than one complete and independent causal explanation of any one event; that is, it does not entail EE. This would only follow if one were to advance the further claim that there is a one-to-one correspondence relationship between a cause and a causal explanation. That is, for explanatory realism and CO to entail EE there must be a one-to-one correspondence relationship between property-instantiations and propositions. And for this to be the case there must be a one-to-one correspondence relationship between properties and predicates. If contrary to this, corresponding to a single property there are various non-synonymous predicates, then from the fact that there can be no more than one cause of a single event it does not follow that there can be no more than one causal explanation of a single event. This is because a single property-instantiation will have various non-equivalent descriptions, and hence corresponding to a single cause there could be complete and independent causal explanations.

Hence whether one considers EE to be correct, ultimately depends upon one's analysis of properties, for whether there is a one-to-one correspondence relationship between predicates and properties depends upon what the criterion of property existence and identity is. It is only given a semantical analysis of properties that such a relationship between predicates and properties can be assumed, and hence that explanations can be individuated in terms of property-instantiations, as EE requires. If a semantical analysis of properties is incorrect then there need not be a one-to-one correspondence of predicates

and properties. Hence proposition E1 may causally explain proposition E3 in virtue of the fact that event e1 bears causal relation R to event e3, but equally proposition E2 may causally explain proposition E3 in virtue of the fact that event e1 bears causal relation R to event e3, where E1 is non-identical with and independent from E2. This is because although the propositions are not identical, they both hold of event e1 in virtue of the same property of e1. To a single property-type there can correspond a multiplicity of non-equivalent predicates, and hence to a single cause there can correspond a multiplicity of complete and non-equivalent causal explanations. Hence a proper understanding of the relationship between predicates and properties makes EE wholly implausible, for in the absence of a semantical criterion, the fact that there are different causal explanations of an event does not imply that there are different causes of that event.

Of course, the possibility remains that Kim does not consider a causal explanation to be a relationship between linguistic entities at all. But if by a 'causal explanation' Kim is referring to a relationship between ontological entities, then he is not really talking about an explanatory relationship at all. It would in fact suggest that by a 'causal explanation' Kim really means a 'cause'. By itself, using 'causal explanation' to mean 'cause' although misleading, is unobjectionable. The problem arises if it is suggested that the Problem of Explanatory Exclusion is anything other than the Problem of Causal Overdetermination, for if a 'causal explanation' just is a 'cause', then obviously the Principle of Explanatory Exclusion is nothing other than the denial of systematic causal overdetermination.

5

A Causal Analysis of Characterising Particulars

5.1 A Causal Criterion of Characterising Particulars' Identity and Existence

If a CP does not depend on its semantic role, either for its identity or for its existence, then upon what does it depend? We have established that there is an intimate link between causation and CPs. For a cause or effect to exist a substance must instantiate a CP, and for C1 to be the same cause as C2, they must be instantiations of the same CP. However, in order to provide an analysis of the existence and the qualitative resemblance of CPs, some philosophers would argue that one must appeal to causal considerations.

In its strongest form, a causal criterion of property existence states that:

E5: P1 exists if and only if P1 makes a causal difference

Referred to by Kim (1993g) as 'Alexander's Dictum', E5 maintains that for a property to be real is for it to bestow causal powers, either upon the single substance that instantiates it (given a theory of CPs), or the various substances that instantiate it (given a theory of universals).²⁸ For example, the CP that is the apple's redness exists if and only if it bestows on the apple causal powers.²⁹

A causal criterion of identity can also be advanced. In terms of universals;

²⁸ For Alexander's formulation see Alexander (1920, p. 8).

²⁹ For this criterion of property existence to be plausible it must be restricted to certain domains of property. For example, it only offers a criterion of existence for those properties instantiated by temporally located substances, for despite the fact that abstract properties such as 'the primeness of the number three' have no causes or effects, one would not want to deny their existence.

- I5: U1 is identical with U1 if and only if they make identical causal differences in all possible circumstances.

In terms of CPs,

- I5*: CP1 exactly resembles CP2 if and only if they make exactly resembling causal differences in all possible circumstances.³⁰

Given I5* one cannot have a single set of exactly resembling CPs, whose members have non-resembling causal powers. If CP1 and CP2 belong to the same exact resemblance class and are instantiated by the same substance (and hence are identical), they will therefore make identical causal differences in all possible circumstances.

Note, that two CPs belong to the same exact resemblance class is not guaranteed by the fact that in a particular circumstance they make exactly resembling causal differences to the substance (or substances) that instantiate them. CP1 and CP2 may be from distinct resemblance classes, and yet 'S1 instantiating CP1' and 'S2 instantiating CP2' may have exactly resembling effects in a particular circumstance. Hence, the heat of this knife tip and the cold of the tip of this ice-cube may give rise to exactly resembling sensations when placed on the back of my neck. However, the two CPs do not have exactly resembling effects in all possible circumstances. For example, they have non-resembling effects when placed in water. For this reason, given I5* they do not belong to the same exact resemblance class.

Equally, as in nearly all cases, a cause is a substance's instantiation of a complex of properties, the causal differences that a property makes, are normally conditional upon the presence of other properties. Hence, to give Shoemaker's (1984b, pp. 211-214) example, if a substance falls under both the type 'being knife-shaped' and 'being made of steel' then it will have the causal power to cut. But all substances that are knife-shaped do

³⁰ Note, this is not to offer an informative criterion of qualitative identity, for it is quite clearly circular, given my understanding of the individuation of the causal relata. A grasp of property identity is needed in order to individuate causes, but causes are being appealed to to individuate properties. Indeed, I do not think that one can offer a reductive analysis of properties. CPs and causal powers belong to a system of internally related concepts and cannot be explained independently of one another. Although there is a fact of the matter about whether CP1 is identical with CP2, there is no non-trivial and non-circular way of fully stating these identity conditions. Hence, rather than a non-circular analysis, this should therefore be seen as offering a non-duplication principle for properties. CP1 is the same as CP2 just in case CP1 and CP2 bestow the same causal powers on the substance that they are instantiated by.

not have the power to cut, for example if they are made of foam. Hence, the causal power of cutting does not necessarily belong to things that are knife-shaped. A knife shaped object has the power of cutting 'conditionally' upon it being made of steel. Hence, for two CPs to be exactly resembling, they must always bestow exactly resembling conditional causal powers upon the substances that instantiate them.

5.2 Applications within the Mental Causation Debate

Clearly, E5 entails the rejection of those properties that are wholly epiphenomenal. If mental properties lack causal powers this in fact leads to their elimination, for given E5 mental realism entails the reality of mental causation. Equally, given I5/ I5*, for mental properties to be distinct from physical properties, the causal powers that they have must not just be the causal powers that physical properties have. One might therefore assume that one could replace the premise of psychophysical causation within the causal closure argument with the premise that mental properties exist. If mental properties exist, the physical world is causally closed, and there is no systematic causal overdetermination, then mental properties must be identical with physical properties. However, this does not follow, for in order to be consistent with the causal closure principle all that must be denied is that non-physical properties make a causal difference within the physical domain. The causal closure principle does not itself entail that non-physical properties cannot make a causal difference within non-physical domains. Thus, for example, it is not inconsistent with parallelism.

But, if Kim (1992, 1993, 1993g) is correct, given this analysis of properties, what must go is non-reductive physicalism. Furthermore, it is not just those particular forms of non-reductive physicalism that appeal to a property layering within a single object that should be rejected. A causal analysis brings into question all forms of non-reductive physicalism, due to the combination of premises that the non-reductive physicalist wishes to maintain. It is a consequence of the non-reductive physicalist's commitment to four premises: That the physical domain is causally closed (contrary to some forms of emergentism), that mental properties exist (contrary to eliminativism), that mental properties are distinct from physical properties (contrary to psychophysical reductionism), and that mental properties supervene upon physical properties.

It is not at all obvious how the non-reductive physicalist can resolve the problem of mental causation, for it is a problem that remains no matter how intimate the connection, short of identity, between mental and physical properties. Mental properties may have their own wholly independent causal powers or they may have causal powers in virtue of the fact that they are identical with physical properties.

On the one hand, the non-reductive physicalist denies that mental properties are identical with physical properties. Hence, if mental properties have causal powers they are not just the causal powers of physical properties. This is because, given I5/ I5*, a causal reduction entails an ontological reduction. Hence, given their commitment to the irreducibility of mental properties, mental properties must make a causal difference over and above that which is made by physical properties. As Kim explains 'To be real, Alexander has said, is to have causal powers; *to be real, new, and irreducible, therefore, must be to have new, irreducible causal powers*' (1993g, p. 350).³¹

But like all physicalists, the non-reductive physicalist must be committed to the causal closure of the physical world. This arguably entails that the causal efficacy of mental properties within the physical domain can be nothing over and above that of the physical. But furthermore, as Kim (1993g) argues, given the relationship of supervenience that the non-reductive physicalist considers to exist between mental and physical properties, mental properties must be wholly epiphenomenal. This is because if mental properties do supervene upon physical properties, it seems that one will always have a subvenient physical property to account for any causal difference supervenient mental properties might be thought to make. So whenever a mental property is appealed to, to causally account for any effect — be it physical or mental, there will be an alternative subvenient physical property which will presumably be sufficient for this effect. Given the denial of systematic causal-overdetermination, this leaves mental properties causally redundant. But if mental properties are wholly epiphenomenal, then, given E5, they do not exist.

If a causal reduction does indeed lead to an ontological reduction, this means that non-reductive physicalism is an unstable position, for its premises are mutually incompatible. One must either reject the causal closure principle, or the claim that mental properties are real, or deny that mental properties are irreducible to physical properties. Hence non-

³¹ Note, that in order to advance this claim, Kim requires not only E5 (That is to be real, is to have new causal powers). In order to maintain that to be irreducible, is to have irreducible causal powers, Kim also needs I5.

reductive physicalism collapses into either interactive mentalism, eliminativism, or psychophysical reductionism.

At this stage, the non-reductive physicalist might respond that his premises are not mutually incompatible. Mental properties make a causal difference, but not in the same kind of way that physical properties make a causal difference, and as mental and physical causes do not causally compete, this is consistent with the causal closure principle. That is he may, in Crane's (1995) words, deny the homogeneity of causation. As Crane explains, that causation is homogenous is an assumption within the argument from causal overdetermination that is commonly overlooked. By the homogeneity of mental and physical causation, Crane explains that he means that the ' "mental" and "physical" as applied to causation are really transferred epithets — what is mental and physical are the relata of causation, not the causation itself' (1995, p.219). Given the homogeneity of causation, mental and physical causation differ only in what they relate — they do not differ in kind.³²

Hence, consider the argument from causal overdetermination. Let us refer to all psychophysical causation as y-causation and all physical causation as x-causation, where x-causation and y-causation differ not only in what they relate, but also differ in kind. It may be true that the physical world is causally-x closed, that is that 'Every physical effect has a set of physical x-causes which together are sufficient for its occurrence' and also true that 'There is y-causation or in other words psychophysical causation'. However, given that x-causation is not y-causation, there is no threat of systematic causal overdetermination, and therefore no need to identify mental causes with physical causes.

³² Jackson (1998) offers an alternative way to respond to the problem of mental causation facing the non-reductive physicalist. This is to appeal to the idea that the physical metaphysically necessitates the mental, and then combine this with a counterfactual theory of causation. In this way one can allow that a mental cause and its subvenient physical base are both causes of a physical effect, without admitting systematic causal overdetermination. Although this response is internally consistent, I find it unpersuasive because it rests upon a theory of causation which I consider to be severely flawed. (See Part 1, §4.6 & Part 3, §6.4) The sheer number of ad hoc postulates that have to be tacked onto a counterfactual theory of causation in order to deal with various counterexamples, should lead us to recognise that counterfactual dependency cannot be all that there is to causation. Consequently, showing that there is a relation of counterfactual dependence between a mental state and a physical effect, is not enough to establish that the mental is causally efficacious within the physical domain.

Here let us return to Yablo's non-reductive physicalist account, an account which, as Crane explains, tries to resolve the problem of mental causation in this kind of way. Even though Yablo accepts the causal closure principle, he denies that all causation is physical causation. His response hinges upon the causal status of a determinate with regard to an effect, when the determinate of this determinable is causally sufficient for the effect. According to Yablo (1992), the causal relevance of a determinate does not preclude the causal relevance of its determinable, and hence determinates and determinables do not compete for causal relevance, even if determinates are causally sufficient for an effect. To give Yablo's example; a pigeon pecks at all and only red objects (1992, p. 257). Of her pecking at a red triangle, one would want to say that it is the redness that is causally relevant. But equally, the particular shade of the triangle — its being scarlet — was causally sufficient for the pecking. Yablo concludes that determinates and determinables are not causal rivals. Hence, despite the causal closure principle and thus the acceptance of the claim that physical determinates are causally sufficient for effects within the physical domain, this does not preclude the possibility that mental properties are causally relevant within the physical domain, because being related as determinable to determinate, mental and physical properties are not causal rivals.³³

However, I sympathise with Crane's view that if non-reductive physicalism responds to the problem of mental causation by denying the homogeneity of causation it 'has lost sight of its motivation' (1995, p. 235). The main motivation for attempting to articulate a supervenience relationship which expresses a dependence relationship between mental and physical properties, was that given his denial of psychophysical reductionism and his acceptance of the causal closure principle, the non-reductionist assumed that he must establish the causal efficacy of mental properties via physical properties. But given the denial of the homogeneity of causation, one can see that there need not be a conflict between mental causation and the causal closure principle. So rather than a non-reductive physicalism why not maintain a full-blooded property mentalism, which accepts the causal closure principle, but denies the homogeneity of causation?³⁴ To this, the non-reductive physicalist will respond that the fact that mental and physical causes do not causally complete, can be best explained if mental properties depend upon physical

³³ Note that Yablo does not make the stronger claim that the causal relevance of a determinate entails the causal relevance of its determinable, and hence that determinables inherit their causal relevance from the relevance of their determinates. (1992, p. 260) This alternative would allow that as mental properties are determinables to physical properties, they inherit their causal relevance from physical properties.

³⁴ Whether the interactive mentalist can advance such a position will be discussed in Part 3.

properties. But I would question this, for I am far from convinced that such non-reductive accounts really do secure the causal efficacy of the mental.

Let us once again return to Yablo's account to see why I think there is a problem. Yablo's denial of the homogeneity of causation cannot be detached from his non-reductive physicalism, for mental and physical properties supposedly do not causally compete, precisely because they are related as determinable to determinate. But do determinables really have any causal powers of their own, or do all the causal powers actually reside in their ultimate determinates? (Equally, one could ask whether second-order properties have any causal powers of their own, or whether all the causal powers reside in the first-order properties that realise them).

My worry is that Yablo's account is not a true denial of the homogeneity of causation. Here we may compare Yablo's account with Kim's. Kim distinguishes between supervenient and subvenient causation. Supervenient causation is set forth as follows: '...when a mental event M causes a physical event P, this is because M is supervenient upon a physical event, P*, and P* causes P....Similarly when mental event M causes another mental event M*, this is so because M supervenes on a physical state P, and similarly M* on P*, and P causes P*' (1993d, p. 106).³⁵ However, in distinguishing between supervenient and subvenient causation, it is questionable whether Kim is really denying the homogeneity of causation because, for Kim, supervenient causation is no more than 'epiphenomenal causation' (p. 106). From this it follows that, although supervenient causation is real and hence 'does not reduce mental causation to the status of a mere chimera' (p. 107), supervenient causation is reducible to subvenient causation (p. 107). Hence, mental causation is reducible to causation taking place at a more basic physical level. From this fact, given I5/ I5*, it follows that mental properties must be reducible to physical properties.

Yablo clearly wants to deny that mental causation is merely 'epiphenomenal causation', but I am not sure that he can, given his understanding of the relationship between the mental and the physical. Indeed, more worryingly, given a proper understanding of determinables, I think that Yablo's account of the mental as a determinable actually entails that there is no such thing as mental causation, and thus given E5, eliminativism.

³⁵ Note that in (1993d) Kim embeds his discussion of supervenient causation within a discussion of macrocausation, and hence is assuming a mereological supervenience rather than a co-instantive supervenience. However, the same points would apply if Kim were assuming a co-instantive supervenience.

To see why I have this worry, let us consider the debate between Fales (1990) and Armstrong (1997) regarding the status of determinables. Yablo is not the only one to maintain that determinables do have causal powers, that are, furthermore, distinct from the causal powers of their determinates. Evan Fales (1990) considers that determinables are universal properties that exist over and above determinates, and defends this claim by appealing to a causal criterion of property identity according to which if two universals stand in different causal relations, then they are distinct, hence a criterion of property identity that is similar to I5.

According to Fales (1990) each of the determinates that falls under a particular determinable will stand in its own distinct causal relations. But furthermore, the determinable will also have a distinct set of causal relations associated with it, that is a subset of each of the sets of causal powers associated with each determinate. Hence, as determinates and determinables have distinct causal powers, this means that they must be distinct. (Fales (1990, p. 239))

Let me give an example to explain what Fales means. Consider the determinable red, and the determinates scarlet and crimson.³⁶ Pigeon 1 pecks at all and only red objects. Call the effect of pigeon 1 pecking e1. Pigeon 2 pecks at all and only scarlet objects. (Call this effect e2). Pigeon 3 pecks at all and only crimson objects. (Call this e3). s1's being a scarlet triangle has effect e1 and e2, but not e3. s2's being a crimson triangle has effect e1 and e3, but not e2. Given the criterion of existence E5, we can conclude that the determinable of being red exists because it has effect e1. Given I5, we can conclude that the determinate scarlet is not identical with the determinable red, because the former has effect e1 and e2, while the latter only has effect e1. And similar reasoning leads us to conclude that the determinate crimson is not identical with the determinable red. Finally, in virtue of the fact that s1 and s2 share a subset of causal relations, namely e1, both fall under the same determinable, namely 'being red'. Hence, we can conclude that determinables exist, and are distinct from determinates.

Now there are various objections to Fales' claims. Armstrong (1997, p. 50) appeals to the criterion of similarity for property identity to reject the idea that determinables exist over and above their determinates. According to this account, if two substances instantiate the same universal (or instantiate CPs from the same exact resemblance class) then there

³⁶ With Fales, I shall assume that properties are universals, although the arguments are equally applicable to sets of exactly resembling CPs.

must be some genuine similarity between them.³⁷ For this reason we should, for example, question the existence of disjunctive and negative universals, universals that a semantical criterion would allow one to accept.³⁸ But it also brings into question the status of determinables. Although two substances that fall under the same ultimate determinate (for example, two objects that are of an identical shade) clearly have something in common, it is not clear that two substances that fall under the same determinable are genuinely similar in virtue of this fact. For example, a crimson object and a scarlet object both fall under the predicate 'x is red', but it is not apparent that we should conclude that there is some respect in which they are ontologically similar in virtue of being red.

One could respond to Armstrong's criticism by enquiring about his interpretation of 'similarity.' For an appeal to similarity to generate an interesting account of property identity, one needs to explain what it means for one property to be genuinely similar to another. Now one plausible response is that for one property to be similar to another they must have similar effects.³⁹ Fales can then respond that, contrary to Armstrong, there is indeed a respect in which substances that instantiate the same determinable universal are similar. In the above example, s1 and s2 are both ontologically similar in virtue of being red, because both have effect e1 in virtue of being red.

But I would suggest that one should be suspicious of Fales' claim that determinables do in fact have causal powers over and above their determinates. Surely the causal powers that the universal 'redness' brings to s1 are nothing over and above the causal powers that the universal 'scarlet' brings to s1, and the causal powers that the universal 'redness' brings to s2 are nothing over and above the causal powers that the universal 'crimson' brings to s2. And, hence, with regard to the pigeon that pecks at all and only red objects, the cause of its pecking at s1 is s1's being scarlet, and the cause of it pecking at s2 is s2's being

³⁷ In terms of CPs, if two substances instantiate CPs from the same set of exactly resembling cp, then there must be some genuine qualitative similarity (resemblance) between them.

³⁸ See Armstrong (1997, pp.26-7) who advances this argument against the existence of disjunctive and negative properties.

³⁹ To this it may be objected that we recognise the similarity of phenomenological properties directly. We do not need to investigate their causal relations with other properties. For example, the fact that crimson is more similar to scarlet than aquamarine is something we recognise without having to consider how their positions differ in the causal network. To this one might appeal to the connection between the intrinsic nature of a property and the causal relations that it enters into. The fact that two properties have the same position in the causal network reflects the fact that they have the same intrinsic nature. In most cases, the exception being phenomenological properties, the only evidence that we have for the sameness of their intrinsic nature is the sameness of their causal relations.

crimson. It is true that the pigeon would have pecked at s1, regardless of whether it was scarlet, that it was the fact that it was red which mattered. But although such an appeal is appropriate to a causal explanation of why the pigeon pecked at both s1 and s2, it is not clear that it is appropriate to the issue of causation, for surely in the particular case of s1 it was s1's being scarlet that caused e1.

To further this argument, let us compare the causal powers of determinables with the causal powers of disjunctives. Let us say that a pigeon pecks at objects if and only if they are scarlet or triangular. s1 is scarlet and circular. s2 is blue and triangular. Hence s1 and s2 both fall under the predicate of being 'scarlet or triangular'. In virtue of this fact, the pigeon will peck at s1 and s2. Because of this causal similarity, should we therefore say that there is a property that they share? Must we be forced to admit that disjunctive properties exist over and above their disjuncts? Surely not, for the pigeon pecks at s1 because it is scarlet and it pecks at s2 because it is triangular. s1's falling under the predicate 'being scarlet or triangular' bring no causal powers to it over and above the ones that its being scarlet does, and s2's falling under the predicate 'being scarlet or triangular' brings no causal powers to it over and above the ones that its being triangular does. It is true that we might say; 'It was the object 'being scarlet or triangular' that caused the pigeon to peck'. But although our causal explanation might appeal to disjunctive predicates, in actual fact it was one of the disjuncts that figured in the causal relation. I would suggest that just as it is implausible that a disjunctive brings new causal powers to a substance that instantiates one of its disjuncts, it is implausible that a determinable brings new causal powers to a substance that instantiates one of its determinates.

Now far from suggesting that determinable properties reduce to determinate properties, with Armstrong, I would want to deny that there are any determinable properties. There are determinable predicates such as 'being red' and substances fall under determinable predicates if they instantiate a determinate property. That is, the having of this determinate property entails that the particular falls under the corresponding determinable predicate (and hence determinable predicates supervene upon determinate predicates). If Armstrong is correct, substances fall under determinable predicates in virtue of nothing more than that determinate universals form a certain class of universals held together by partial identities.⁴⁰ Hence, 'S's being red caused the pigeon to peck' is true not because there is a universal 'being red' the instantiation of which caused the pigeon to peck.

⁴⁰ See Armstrong (1997, pp.50-52).

Rather, S exemplifies some fully determinate universal, for example, 'being scarlet' which is a member of a certain class.

Given this analysis of determinables, Yablo is indeed correct to suggest that determinates and determinables are not causal rivals, but this will not be because the causal relations that determinables enter into do not compete with those that determinates enter into. Rather, as there are no determinable properties, there is no determinable causation. If mental predicates are determinable predicates, there are no mental properties and hence there is no mental causation. 's's being in pain caused e' is true, not because there is a mental property, the instantiation of which by s caused e, but because s exemplifies some fully determinate physical property which is a member of a certain determinable class. If this is the case, the causal relevance that Yablo considers mental determinables to have can at best point to the fact that there are causal explanations that cannot be picked out using purely physical determinate predicates. And, of course, the fact that we appeal to mental determinables in causal explanations does nothing to show the reality of mental causation. For given the rejection of a semantical criterion of property analysis, we can see that the fact that determinate and determinable predicates arise at different explanatory levels, is consistent with substances falling under determinable predicates in virtue of instantiating non-determinable (determinate) properties.

Because of the relation that Yablo considers the mental and physical to have, his attempt to deny the homogeneity of causation in order to avoid epiphenomenalism, is not a true denial of the homogeneity of causation. Yablo's appeal to the causal relevance of the mental instead merely points to the fact that mental predicates play a non-redundant role in causal explanations of physical states, and his non-reductive physicalism is actually a disguised eliminativism. Do these considerations about Yablo's non-reductive physicalism generalise to all forms of non-reductive physicalism that attempt to respond to the problem of mental causation by denying the homogeneity of causation?

Well, to be fair, I think that one needs to examine the various attempts to deny the homogeneity of causation individually, considering the particular relationship that the non-reductive physicalist considers to exist between mental and physical properties, and the particular way in which he attempts to deny the homogeneity of causation. However, the general considerations that have led to the rejection of Yablo's account, should make us highly suspicious of all of these kinds of physicalist account, and in particular, those that base their non-reductive physicalism within a co-instantive supervenience account. Heil and Robb have drawn our attention to the fact that the rejection of a semantical analysis

of CPs means that we should be suspicious that there is any property layering within an object. And this should not be forgotten when we consider such non-reductive physicalist's attempts to deny the homogeneity of causation and hence ascribe the mental with causal efficacy. Rather than denying the homogeneity of causation, are they merely pointing to the fact that mental and physical predicates arise at different levels of explanation? Given the rejection of a semantical account of property existence and identity, this should not lead us to conclude that mental and physical properties arise at different ontological levels.

To summarise the points that have been made about non-reductive physicalism. We started by considering Heil and Robb's claim that a rejection of a semantical criterion leads to the rejection of all property layers within an object. This would entail the rejection of all those forms of non-reductive physicalism based within a co-instantive supervenience. The point was then raised that, contrary to Heil and Robb's suggestion, all that the rejection of a semantical criterion shows is that, from the fact that there are different levels of explanation, one cannot conclude that there are different ontological levels of properties. It does not establish that there are not any property-levels within an object. To show this one must advance a positive analysis of CP identity.

We therefore turned to a strong causal criterion of property existence and identity. Given this criterion, the non-reductive physicalist's position is an untenable one. This is because, given their commitment to the causal closure principle and psychophysical supervenience, unless one denies systematic causal overdetermination, mental properties must have the status of mere epiphenomena. And as, according to the causal criterion, a causal reduction leads to an ontological one, contrary to the non-reductive physicalists' claims, mental properties cannot be anything over and above their subvenient bases. However, many non-reductive physicalists try to get around this problem by denying the homogeneity of causation. But do such accounts really manage to secure the causal efficacy of the mental? Taking Yablo's theory of mental causation as my central example of such an account, I argued that far from providing mental properties with causal efficacy, his 'denial of the homogeneity of causation' in fact only establishes that mental predicates play a non-redundant role within causal explanation. And as explanatory levels do not lead to the existence of ontological levels this does nothing to establish that there is mental causation. In fact, a proper understanding of the relationship between determinables and determinates suggests that if mental predicates are determinables, then there is no mental causation, because there will be no mental properties. Our discussion of Yablo suggests that we should be suspicious of any non-reductive

physicalism that attempts to avoid epiphenomenalism by denying the homogeneity of causation.

A proper understanding of a CP's existence and identity threatens the plausibility of non-reductive physicalism, but equally, its motivation has been removed for the rejection of a semantical analysis leads to the rejection of the multiple realisability argument. Consequently, a psychophysical reductionism becomes far more plausible. But here let us make two observations about psychophysical reductionism. Heil and Robb adopt a psychophysical reductionism. They appear to assume that if there are no property layers within an object, then this collapses the position of those philosophers of mind who consider mental and physical properties to occupy different ontological levels within a single object into psychophysical reductionism. However, if, for example, the mental and physical are related as determinable to determinate, far from yielding a psychophysical reductionism, it arguably forces one to accept an eliminativism. Heil and Robb dismiss eliminativism as 'at odds with manifest experience' (forthcoming, p. 1). But the worry is that if they are to stay true to their aim to provide a theory of mind that is ontologically driven, they need to do more to distance their position within the mental causation debate from eliminativism. In order to show that mental CPs are identical with physical CPs, they must first show that mental CPs exist.

Secondly, obviously the rejection of the multiple realisability argument does not serve to establish psychophysical reductionism. Which form of physicalism to accept is an in-house debate. For the full-blooded interactive mentalist who thinks that mental causes have independent effects within the physical domain, either because they reject the causal closure principle, the homogeneity of causation or the premise of the denial of systematic overdetermination, physicalism is no more attractive. What property analysis has enabled us to do is to thin out the physicalist positions and hence decide who the true contenders within the mental causation debate are. The crucial question then is whether, if mental CPs are not identical with physical CPs, can mental causes have real independent effects within the physical domain? This is the question that Part Three is concerned with. The final part of this section will consider a different kind of property layering — one between different objects.

6

Mereology

This discussion has focused on whether a single object can instantiate properties that are hierarchically ordered. But there is the further question of whether there is a hierarchical layering of properties between distinct objects. Segal and Sober (1991), and Kim (1993l), having rejected co-instantive supervenience as a viable theory of the mind-body relation, suggest that we should understand psychophysical supervenience as an instance of mereological supervenience. According to it, the properties and relations of a whole, supervene upon the properties and relations of its parts. That is, two macrophysical properties cannot differ without some difference in the properties of, or the relations between, their microphysical components.

What is the ontological significance of mereological supervenience? Well, non-reductive physicalists who appeal to mereological supervenience, assume that it leads to the existence of levels of properties. The world is layered in a single hierarchical structure from parts to wholes. At the bottom level are entities that are most basic. At any given level, except the bottom level, each entity within it has a complete decomposition into entities belonging to the level below it. New properties arise at the different levels on the macro-micro hierarchy, and the non-reductive physicalist considers the relationship between these different levels of properties to be one of supervenience.

The non-reductive physicalist, who appeals to a relation of mereological supervenience between mental and physical properties, will consider his account to have an immediate advantage over those forms of non-reductive physicalism that appeal to a relationship of co-instantive supervenience between mental and physical properties. This is because, while the latter struggled to establish a dependency relationship between the mental and the physical, mereological supervenience is structured by the part-whole relationship into which a dependency relationship is built.

It should be observed, that there has been a great amount of confusion between co-instantive supervenience and mereological supervenience, and the tendency has been to simply switch from one to the other. Hence, for example, in 'The nonreductivist's troubles with mental causation', Kim (1993g) confuses the two different kinds of property layering

that these accounts of supervenience appeal to. He begins by discussing a non-reductive physicalism that assumes that mental properties and physical properties are instantiated by different substances on the micro-macro hierarchy. (p. 337) He then goes on to discuss a 'basic tenet of non-reductive physicalism'; that 'for a mental property to be instantiated in a system, that system must instantiate an appropriate physical property...' (p. 347) (*My emphasis*). This is to assume a non-reductive physicalism based within a co-instantive supervenience account.⁴¹ But one must distinguish between the different types of property layering that these different kinds of non-reductive physicalism appeal to. A non-reductive physicalist who assumes a co-instantive supervenience is appealing to a property layering within a single object. A non-reductive physicalist who assumes a mereological supervenience is appealing to a layering of properties between distinct objects.

For this reason, one cannot simply take a theory of psychophysical co-instantive supervenience, and without any adjustment, project it onto a layered model of the world in order to advance a mereological supervenience. How then should the non-reductive physicalist fit mental properties onto the layered model of the world, for the objects that instantiate mental properties must be different from the objects whose properties they supervene upon? One possibility is that mental properties arise when the brain is taken as a holistic whole. Individual neurons have electrical properties but do not possess mental properties, and when neurons are combined to form a system, mental properties arise. The non-reductive physicalist would then maintain that mental properties are supervenient upon the properties of, and relations between, the individual neurons. That is to say, there can be no change in the properties of the neurological system without there being a change in the properties of the neurons that compose it.

Having seen how the non-reductive physicalist might base his theory within a mereological supervenience account, the important question is whether an appeal to mereological supervenience actually allows one to advance a plausible form of non-reductive physicalism. The non-reductive physicalist requires that mental properties are distinct from physical properties, and that the former supervene upon the latter, where supervenience is not mere covariance. Rather, properties of a whole must in some way depend upon the properties of their parts. To see whether the properties of a whole and the properties of its parts really do exhibit this kind of relationship, and if so, whether such a dependency relationship establishes non-reductive physicalism, I shall start by

⁴¹ Also see Robb (1997, p. 185). Robb assumes a co-instantiation theory, but then raises the question of whether the problem of mental causation carries over to the special sciences.

considering whether there is a dependency relationship between a complex object or whole and the various objects that compose it, and if so, of what kind. This will then lead us on to the question of whether there is a dependence relationship between the properties of a whole and the properties of its parts.

Complex objects have objects as parts; they are composed or constituted by their constituent objects. The relationship between a complex object, or, in other words a whole, and its parts is one of dependence, although the dependence relationship differs in its degree of strictness depending upon the whole under consideration.⁴² Some wholes, namely sums, strictly depend upon their parts. A sum exists purely in virtue of its parts existing and thus cannot persist identically through a mereological change. Hence, for example, a pile of stones is a sum. If a stone is removed from it, although we still have a collection of stones, we no longer have the same collection. But sums are the exception, for most wholes do not strictly depend upon their parts. Unless one maintains mereological essentialism, given the rejection of a doctrine of temporal parts, one would want to argue that most objects are mereologically variable. Unlike a sum, compound material objects and organisms may, within certain limits, gain, lose, or alter their parts without ceasing to exist. For example, provided it does not cause its collapse, if a couple of stones that belong to a wall are destroyed, the collection of stones ceases to exist, but the wall does not. Equally, the wall could have a number of its stones replaced, and continue to exist. Hence the wall is not a mereological sum of the stones. As it can survive a change in its parts, the relationship between the wall and the stones that compose it is not one of strict dependence.⁴³

However, there is a weaker dependency relationship between such a whole and its parts. The wall could not exist through any change whatever in its parts or the relations between them; its form must be preserved. Hence, if the stones that constitute the wall were to be scattered, or if many of the stones were destroyed, then the wall would cease to exist. There is, in other words, a kind of 'generic' dependence between such a whole and the parts that constitute it. Here, we may appeal to Lowe's (1998) definition of generic

⁴² I am assuming that there are continuants. That is, there are objects that exist in time, come into existence, persist, and then cease to exist. Contrary to the doctrine of temporal parts, if a continuant has contrary properties at different times it is the whole continuant and not just part of it, which has different properties successively.

⁴³ Returning to Lowe's (1998) definition of ontological dependence, such an object does not ontologically depend upon its proper parts, as the identity of the object is not wholly dependent upon the identity of its parts. (See Part Two, §3.3).

dependence. According to Lowe, 'x depends for its existence upon objects of type T if necessarily, x exists only if something y exists such that y is of type T' (1998, p. 141). This is clearly the kind of dependency relationship that we should be concerned with in the mental causation debate. If, for example, we consider mental properties to be properties of a neurological system, then this neurological system is clearly not just a sum of the neurons from which it is composed. The neurological system may persist through certain changes in its parts. Also note that the idea that there is a generic dependency between a whole and its parts is consistent with there being a relationship of supervenience between them. That is, it is consistent with the claim that there can be no change in the whole without there being a change in the mereological components.

But is there a corresponding dependency relationship between the properties of a whole and the properties of its parts? Well, arguably in all cases the properties of a whole are a product of the properties of their parts, when account is taken of the arrangement of their parts, together with the various causal interactions between these parts.⁴⁴ This is most obviously the case with properties such as volume and length. These kinds of property may be a property of both a whole and its parts, although the whole may have a different value of the property to that of its parts. For example, the volume of a building has a value which none of its parts have, although they each possess a volume. These kinds of property do exhibit a dependency. The building having the volume and height that it has, is wholly dependent upon the properties and organisation of its parts. (Of course, it is not strictly dependent upon these very parts; one brick could be replaced by another with the same volume, resulting in the same net volume).

New properties arise at different levels within the macro-micro hierarchy. For example, at the cellular level metabolism arises. The property of metabolism is not ascribable to an entity below the cellular level, for example a molecule. However, a dependency relationship between these properties is still maintained, for the existence of these properties can be explained when one takes into account the various interrelations and interactions between the properties of the parts that compose cellular objects. The same, one could argue, is true of mental properties. The existence of mental properties depends upon the existence of, and the interrelations and the interactions between, the properties of those physical objects that compose mental objects.

⁴⁴ An arguable exception is a quantum mechanical states, such as those of a hydrogen atom. Arguably one could have two quantum mechanical states that were exactly the same in respect of their parts and the relations between their parts but which differed as a global system. See Teller (1992) for a discussion of this issue.

Let us assume that there is such a dependency relationship between mental properties and the properties of the physical objects that compose mental objects. The question that we now need to ask is, if mental properties are a product of, depend for their existence upon, the properties of their physical parts, then what should we conclude from this? Well, it shows that any causal powers that the properties of a mental object has, originate from the causal powers of the properties of its physical parts. That is, that the causal powers of mental properties originate from the causal powers of physical properties. But I would suggest that this is in fact consistent with three conflicting positions within the mental causation debate.

Firstly, the claim that the causal powers of mental properties originate from the causal powers of physical properties is consistent with non-reductive physicalism. In addition to the claim that the causal powers of the properties of a whole originate from the causal powers of the properties of its parts, a non-reductive physicalist maintains that the causal powers of the properties of a whole are distinct from, but entirely determined by or grounded in the causal powers of the properties its parts. Because the properties of a whole are distinct from the properties of a part, this leads to a layering of properties, and hence a layering of objects. Because the causal powers of the properties of a whole are entirely grounded in the causal powers of the properties of the parts from which they are composed, the non-reductive physicalist is able to stay true to physicalism.

For example, consider the causal status that Searle (1992, ch. 5) allows mental properties. According to Searle, there are emergent properties that bring new causal powers with them. Water has the power to dissolve; hydrogen and oxygen do not. A heap of material possesses instability; none of the individual parts do. However, the causal powers that these properties bring can be fully analysed in terms of the causal powers of a things parts, when account is taken of the parts, their arrangement and their causal interactions with one another. According to Searle, consciousness is an emergent property in this sense. Its existence can be explained when account is taken of parts of the brain, their arrangement and the causal interactions between these parts. Like solubility and stability, consciousness does not possess any causal powers that are not causally reducible to the causal powers of their parts. In his claim that mental properties are distinct from physical properties, but have no independent causal powers, Searle

could therefore be interpreted as supporting a non-reductive physicalism.⁴⁵ Certainly, this is the kind of causal account of mental properties that the non-reductive physicalist who bases his theory within a mereological supervenience account would want to advance.

However, just as non-reductive physicalism is an unstable position when based within a co-instantive supervenience account, it is unstable when based within a mereological supervenience account. With the case of a non-reductive physicalism based within a co-instantive supervenience account, we observed that given his acceptance of the causal closure principle and his denial of systematic causal overdetermination, he is arguably forced to say that the causal powers of a mental property are nothing over and above the causal powers of physical properties. But if this is the case, given I5, this entails that mental properties are nothing over and above physical properties. The same considerations equally apply to a non-reductive physicalism that appeals to mereological supervenience. If the causal powers of the properties of a whole are reducible to, or 'nothing over and above' the causal powers of the properties of the parts from which they emerge, then given I5, it follows that the properties of a whole must be reducible to the properties of its parts. A causal analysis of properties leads to the rejection of any position that maintains that there are property levels whilst denying that there is any genuine downward causation (or horizontal causation at the emergent level). If the only genuine causation takes place at the level of physics, then the only genuine properties that there are are physical properties, and hence there will be no property levels. In short, if mental properties arise at a level on the macro-micro hierarchy that is distinct from the level of the properties of their parts, then given the causal analysis of properties, mental properties must make a genuine causal difference over and above that of the properties from which they emerge.

At this stage, it might be objected that this kind of non-reductive physicalism could provide the mental with causal powers by denying the homogeneity of causation, rather in the way that Yablo attempts to. But I do not think that this is an option for the kind of non-reductive physicalism under consideration here. If such a non-reductive physicalist denies the homogeneity of causation, I would suggest that there will be nothing to distinguish his position from an interactive mentalism. Yes, the causal powers of mental properties will

⁴⁵ I am reluctant to press the similarity between Searle and the non-reductive physicalist, for there is controversy over whether Searle is really intending a non-reductive physicalism, for Searle thinks that the irreducibility of consciousness does not arise in virtue of any ontological reason, but trivially from our definition. See Searle (1992, p. 122). For a discussion of this point see Lowe (1996, p. 635). Also see Heil (1992, pp. 125ff) who maintains that Searle is offering a type-identity theory.

originate from the causal powers of physical properties, but from this fact it does not follow that mental properties do not have independent causal powers.⁴⁶ To truly deny the homogeneity of causation will surely therefore be to allow that mental properties have independent causal powers.

Furthermore, note that some forms of emergentism, for example those commonly associated with British emergentism, despite endowing mental properties with independent causal powers do not conflict with mereological supervenience.⁴⁷ Mental properties have causal powers over and above the causal powers of physical properties, but they only operate alongside physical properties. The fact that it is compatible with mereological supervenience that mental properties can be fundamental causal properties, which have causal powers over and above those of physical properties, shows that mereological supervenience is still not enough to entail non-reductive physicalism.

If, on the other hand, mental properties do not have irreducible causal powers this entails, given I5/ I5*, that mental properties do not exist on a different ontological level to the properties of their physical parts, for a causal reduction leads to an ontological one. This is to therefore maintain a psychophysical reductionism. More generally, if there are no emergent properties, that is, if the causal powers of all macro-properties are nothing over and above the causal powers of micro-properties acting together, then one would have to reject the idea that there are any different levels of properties. And obviously, if there are no property levels, there are not different levels of objects. There is, in other words no macro-micro hierarchy.

This is Heil and Martin's (1999) and Heil's (1998b) position. According to them there is no vertical causation. Consequently, there are no layers of properties and objects. There is only a single level of ultimate constituents. Heil and Martin maintain that the properties of a whole are identical with the properties of its parts when account is taken of all its parts and their interrelations and interactions. This is not to suggest that mereological wholes fail to exist. Rather, they just are not ontologically additional to the collection all of their parts, because the causal powers that the properties of a mereological whole has, are nothing more than the causal powers of the properties of all of its parts acting together.

Heil and Martin's account shows that one can identify mental properties with physical properties, despite agreeing with the non-reductive physicalist that mental properties are

⁴⁶ See Lowe (1993, pp. 635-637) for further defence of this point.

⁴⁷ See Horgan (1993, pp. 557- 560) for a discussion of this point.

related to physical properties as the properties of a whole to the properties of its parts.⁴⁸ Furthermore, one can agree that mental properties mereologically supervene upon physical properties, for Heil and Martin's position is quite compatible with the idea that mental properties mereologically supervene upon physical properties. As we have seen if P1 supervenes upon P2, then this is compatible with P1 being identical with P2. Indeed, Armstrong (1997, pp. 11-12), appeals to the notion of 'mereological supervenience' but in doing so means that mereological wholes are identical with all their parts taken together.

One may argue that in rejecting the layered view of reality, Heil and Martin are also rejecting the plausible claim that the various scientific disciplines are concerned with objects and properties occupying distinct ontological levels. However, our rejection of a semantical criterion of property analysis means that the rejection of ontological layers does not lead to the rejection of explanatory layers. Consequently, Heil and Martin can maintain that the different sciences are concerned with different levels of explanation. Chemical explanations, for example, appeal to chemical concepts and predicates, and these cannot be expressed in the language of physics. But although chemical explanations cannot be reduced to physical explanations, given the rejection of a semantical criterion, this does not entail that the ontology of chemistry cannot be reduced to the ontology of physics. Furthermore, from the rejection of Kim's Principle of Explanatory Exclusion, it follows that even if physics is explanatorily adequate, that is, even if every physical occurrence is entirely explicable by reference to comparable physical occurrences, this does not mean that every non-physical explanation of a physical occurrence must be identical with or dependent upon a physical explanation. Hence, even if physics is explanatory adequate, this still leaves room, for example, for independent chemical explanations of physical occurrences.

Heil and Martin's rejection of a layered model of reality is based upon the premise that there is no vertical causation. This premise will be rejected by the final kind of position that would accept that mental properties depend for their existence upon the properties of their physical parts – that is, an emergentism or full-blooded interactive property mentalism. With the non-reductive physicalist, the mentalist agrees that mental properties are emergent properties which arise within a system when it reaches a degree of complexity,

⁴⁸ Heil and Martin's rejection of ontological levels only entails psychophysical reductionism if mental properties are macro-properties. With Heil and Martin one may reject the existence of vertical causation, and hence given a causal analysis of properties, the existence of ontological levels, but advance a full-blooded mentalism which maintains that mental properties are instantiated by simple substances. Whether these simple substances also instantiate physical properties will be a further question.

and hence, although non-physical, are a product of physical processes. Hence, mentalism is wholly consistent with the idea that two objects with an identical total microstructure will instantiate exactly the same macro-properties.

However, unlike the non-reductive physicalist, a full-blooded interactive property mentalism maintains that mental properties have independent causal powers. For although the properties of a whole originate from the properties of its parts, and hence the causal powers of the properties of a whole originate from the causal powers of the properties of its parts, this does not entail that properties of a whole do not have independent causal powers, that is, that their causal powers must be grounded in and determined by the causal powers of their physical parts.

Hence, if mental properties are related to physical properties as the properties of a whole to the properties of its parts, then it is plausible to conclude that the causal powers that mental properties have, originate from the causal powers that physical properties have. This is consistent with three conflicting positions within the mental causation debate. It is consistent with non-reductive physicalism, psychophysical reductionism and interactive property mentalism. Our causal analysis of CPs leads us to reject the first option. Whether one advances a psychophysical reductionism or an interactive mentalism, will be determined by whether one thinks that there is vertical causation. To suggest that emergent mental properties have physical effects commits one to the idea that mental properties could causally affect lower levels, levels from which they emerge, hence entailing the denial of either the causal closure principle or the homogeneity of causation. Equally, there is the further issue, raised by Kim (1999b) and Heil and Martin (1999), of how downwards causation is possible. That is, how higher-level properties could make a causal difference to the properties from which they arise. I would suggest that how well the emergentist is able to address both of these issues, depends upon what he is understanding psychophysical causation to actually consist in, which itself depends upon the theory of causation that he is basing his theory of psychophysical causation in. These issues will be the subject of Part Three.

Part Three

Causal Closure Principles

Causal Closure Arguments

1.1 Preliminaries

The causal relata are states, where states are instantiations of CPs. Given a plausible analysis of CPs, certain positions within the mental causation debate can be rejected. In particular, it is highly implausible that there is any plausible physicalist middle ground. The true physicalist contenders are those that identify mental and physical CPs, and those that eliminate mental CPs. And assuming that mental CPs make a real causal difference within the physical domain, psychophysical reductionism is the only physicalist option. Indeed, as the rejection of a semantical analysis of CPs leads to the rejection of the problem of multiple realisability, one of the main barriers standing in the way of adopting a psychophysical reductionism has been removed.

However, this should not serve to sway the interactive dualist who rejects the identification of mental CPs with physical CPs, not because of considerations of multiple realisability, but because he considers mental causation to be distinct from physical causation, and hence the identification of mental causes with physical causes to be wholly misguided. Interactive mentalists reject the central premise found within the best arguments for psychophysical reductionism — the causal closure principle. Causal closure principles, or more precisely, those causal closure principles that are strong enough to yield a physicalist conclusion, are false.

Hence, the real choice in the mental causation debate is between a psychophysical reductionism and an interactive mentalism, and which is correct depends upon whether a causal closure principle that is strong enough to establish psychophysical reductionism is plausible. The aim of Part Three is to show that it is far from evident that any such causal closure principle is plausible.

1.2 Two Different Kinds of Causal Closure Argument

i) Strong Causal Closure Arguments

Various causal closure principles are appealed to within the mental causation debate. Smith and Jones (1986), for example, understand the causal closure principle to be the principle that 'physical events do not have non-physical causes' (p. 66). On the other hand, both Kim (1993f) and Baker (1993) advance a causal closure principle according to which: 'Any physical event which has a cause at time *t* has a physical cause at time *t*'. Kim considers this to entail that 'if we trace the causal ancestry of a physical event, we need never go outside the physical domain' (1993f, p.280). This latter way of formulating the closure principle is also suggested by Robb (1997, p. 183), LePore and Loewer (1987, p. 630). Hence, according to Robb the closure principle is the principle that 'every physical event has in its causal history only physical events and physical properties' (1997, p.183). Papineau advances a similar form of closure principle according to which 'All physical effects are fully determined by law by a purely physical prior history' (2000, p. 179), but elsewhere interprets the claim as; 'Every physical effect has a sufficient physical cause' (1998, p. 375).

I would suggest that these various formulations of the closure principle can in fact be divided into two main groups; those that simply require the additional premise of psychophysical causation in order to yield a monistic conclusion, and those that also require the further premise of the denial of systematic causation overdetermination.¹

The strongest form of argument for monism that appeals to a causal closure principle consists of merely two premises; the premise that mental causes have (physical) effects, and a causal closure principle, which given the first premise, is strong enough to allow one to conclude that mental causes must be identical with physical causes. Such a principle can be formulated in one of two ways.

Firstly, as the claim that 'No effect has a non-physical cause.' According to it, causation is essentially physical and thus even the possibility of intra-psychological causation is denied. Unless mental CPs are identical with physical CPs, mental CPs will be wholly

¹ Our discussion so far has assumed the latter kind of formulation of the causal closure argument.

epiphenomenal, and thus given a strong causal analysis of CPs, must be eliminated. As this strong form of closure principle yields a monistic conclusion, even if it is combined with the much weaker premise that mental causes have effects, the causal closure argument can be formulated in the following way;

- 1. No effect has a non-physical cause
- 2. Mental causes have effects

Mental causes are identical with physical causes

The second weaker form of causal closure principle maintains only that 'No physical effect has a non-physical cause'. While the acceptance of this premise does not lead directly to the rejection of intra-psychological causation, unless one accepts psychophysical reductionism, one must deny that there is any psychophysical causation. Hence;

- 1. No physical effect has a non-physical cause
- 2. Mental causes have physical effects

Mental causes are identical with physical causes

These two kinds of argument for psychophysical reductionism shall be referred to as 'strong causal closure arguments'. Arguably, the causal closure principles that they appeal to should be rejected by interactive mentalists as question-begging, for such forms of mentalism consist precisely in the claims that: (1) Mental causes are not identical with physical causes and (2) Mental causes do have (physical) effects. I will later consider whether these causal closure principles really are based upon mere monistic prejudice, or whether they in fact have some metaphysical or empirical grounding, but first I want to spend some time analysing a weaker but far more popular kind of causal closure argument.

ii) The Argument from Causal Overdetermination

A much weaker kind of causal closure principle consists roughly in the claim that 'Every physical effect has a sufficient physical cause.' If this closure principle is combined with the premise of psychophysical causation and the further premise that there is no

systematic causal overdetermination, then a monistic conclusion appears to follow. Hence;

1. Every physical effect has a sufficient physical cause
2. Mental causes have physical effects
3. There is no systematic causal overdetermination.

Mental causes are identical with physical causes

Unlike the two previous forms of causal closure argument, none of the individual premises of the argument from causal overdetermination rule out the possibility of dualistic psychophysical causation. As it is consistent with the possibility that mental causes could have physical effects, it therefore cannot be accused of begging the question against interactive mentalism. Its point is not that mental causes cannot have physical effects, but that physics never needs to appeal to such causes in order to give a complete causal account of its effects, because a physical effect that has a mental cause will also have a complete physical cause, hence rendering the mental cause redundant. Equally, as this causal closure principle is considerably weaker than the others, it should be easier to support.

A number of points need to be made about this kind of closure principle. Firstly, it is not claiming that all physical states have sufficient physical causes, and hence is consistent with the possibility of their being uncaused physical states. Secondly, a causal closure principle should allow for the possibility of indeterministic causation. Therefore, one should define the closure principle in the kind of way that Papineau (1993) does: 'A purely physical specification, plus physical laws, will always suffice to tell us what is physically going to happen, insofar as that can be foretold at all' (p.16). However, as none of the arguments that I raise are affected by this consideration, in order to simplify matters, one can talk as if all causation were deterministic. Thirdly, the kind of causal overdetermination that we are concerned with arises when an effect is the product of more than one causal chain which would each have been sufficient to produce the effect. (That is, with independent causal chains). Fourthly, the causal closure principle's concern should be with the complete set of causes that together are causally sufficient for the occurrence of an effect. Hence, the causal closure principle should be understood as the principle that;

'Every physical effect has a set of physical causes which together are sufficient for its occurrence.'

What is to count as the complete set of causes of an effect? Well, I reject the distinction between the cause and the causal field. This is a distinction that Mackie defends within his discussion of causation. For example, although the fact that a person lives, breathes and eats, matters to whether he contracts an illness, Mackie would not regard them as causes of the contraction of the illness. Instead, they are part of the causal field, 'a background against which the causing goes on' (Mackie (1980, p. 63)). If the causal field is not a part of the cause of an effect, then it could be true that although physical causes suffice for physical effects, mental states do play a causal role as they form an essential part of the causal field that is required in order for the physical cause to have its effect. For example, one might argue that a person's desires are part of the background circumstances required in order for a neurological state of his brain to cause his arm to raise. Given the distinction between the cause and the causal field, this would be consistent with the above causal closure claim. This response is not available to me, because what Mackie takes to be the causal field, I take to be real causes that have, for pragmatic reasons, been relegated to the background. Hence, all of these causal factors are in fact part of the complete set of states that together are sufficient for the cause to take place. Thus, I consider the claim that a person's desires are part of the background circumstances needed in order for his arm to move, to be inconsistent with the claim that every physical effect has a complete set of sufficient physical causes.

However, even if the claim that 'Every physical effect has a complete set of sufficient physical causes' is inserted into the argument from causal overdetermination, this is not strong enough to yield monism. Within the argument from causal overdetermination, the role of the causal closure principle is to rule out the possibility of there being any causal 'gaps' in the physical domain, which non-physical causes might fill. It is generally assumed that the redundancy of non-physical causes within the physical domain does follow from something like; 'Every physical effect has a complete set of sufficient physical causes'. For example, Chalmers informs us that 'the physical world is more or less causally closed: for every physical event, there is a physically sufficient cause. If so, there is no room for a mental 'ghost in the machine' to do any extra causal work' (Chalmers (1996, p.125)). Hence, the assumption is that if a cause C1 is causally sufficient for an effect E, then given the denial of systematic causal overdetermination, there is no room for any other cause of E that is distinct from C1.

But this assumption is false, for as Lowe (1999a, p. 229-30 & 2000, p. 6-8) argues, it ignores the transitivity of causation. Given the transitivity of causation, if C1 is causally

sufficient for C2, and C2 is causally sufficient for E, then C1 is causally sufficient for E. Hence both C1 and C2 are causally sufficient for E, but this does not mean that C1 and C2 causally overdetermine E. Hence, let us say that E is a physical effect. If C1 is physical, given the transitivity of causation it follows that the physical effect E has a sufficient physical cause, regardless of whether C2 is a physical cause. That is, if the physical state C1 is causally sufficient for a mental state C2 which in turn is causally sufficient for the physical state E, then it is still true that the physical effect E has a sufficient physical cause because due to the transitivity of causation, C1 is causally sufficient for E. Hence if the closure principle that 'Every physical effect has a complete set of sufficient physical causes' is inserted into the argument from causal overdetermination, far from entailing monism, it is in fact consistent with emergentism, because the causal closure principle is consistent with mental causes figuring in the causal chain between the initial physical cause and the physical effect that it is causally sufficient for. All that is required is that the initial cause within a causal chain involving a physical effect is physical, but this is exactly what emergentism maintains. According to such a position, from physical properties, mental properties have emerged which themselves have independent causal powers within the physical domain.

Can the premise of causal closure be strengthened within the argument from causal overdetermination without making the premise of the denial of systematic causal-overdetermination redundant? That is, is there a form of causal closure argument that entails monism, but which does not reduce to one of the strong causal closure arguments? Such a monism wishes to capture the idea that physics never *needs* to appeal to mental causes in order to give a complete causal account of its effects, because there are no gaps in physical chains of causation which mental causes are required to fill. I would suggest that this could best be done by advancing a causal closure principle that maintains that:

'Every physical effect has a set of direct physical causes which together are sufficient for its occurrence.'

Considerations of the transitivity of causation reveal that causes need not be direct causes. C1 is a cause of E, either if C1 is a direct cause of E, or if there is a chain of direct causation leading from C1 to E. If C1 causes C2 which causes E, although C1 is a cause of E, C1 is not a direct cause of E, because there is a further causal intermediary between C1 and E. It was precisely because the previous causal closure claim was consistent with physical effects merely having indirect physical causes, and hence mental causal

intermediaries, that it was compatible with emergentism. If however, every physical effect has a direct physical cause which is sufficient for its physical effects, there will be no such gaps in the causal chains, that mental causes are needed to fill. Neither does this closure principle dismiss the very possibility of psychophysical causation; mental causes are simply not required to play a causal role in the physical domain. Consequently, if the monist appeals to the argument from causal overdetermination, I would suggest that this is the strength of causal closure principle that he requires.

1.3 How to Assess the Plausibility of a Causal Closure Principle

We are presented with two different types of causal closure principle, which when inserted into a causal closure argument would appear to yield monism. A strong causal closure principle denies either the existence of non-physical causation or the existence of non-physical/ physical causation. A weaker and very different kind of closure principle, namely, that: 'Every physical effect has a set of direct physical causes which together are sufficient for its occurrence' denies only that physics will ever need to appeal to non-physical causes.

How should we assess the plausibility of a causal closure principle? The plausibility and strength of a particular formulation of the causal closure principle is affected by what one understands a physical cause to be, and hence how one defines the term 'physical.' For example, clearly a trivial physicalism that includes within its definition of the term 'physical' those categories such as the mental that one is wishing to contrast the physical with, will generate a causal closure principle that not even the interactive mentalist would wish to reject. The problem for the monist is that of advancing a definition of the 'physical' which is non-trivial, and yet which is plausibly causally closed in the sense that he requires.

Less commonly noticed is that the plausibility of a causal closure principle also depends upon what causation actually is. I will argue that the more broad-minded one's theory of causation, the more difficult it will be to justify a causal closure principle. Certainly, whether certain strong formulations of the causal closure principle can plausibly be allowed depends upon whether certain strong theories of causation can be allowed. But furthermore, what causation is, is also very important for assessing the plausibility of the weaker kind of closure principle appealed to in the argument from causal overdetermination; What it is to be a 'cause' and what it is to be an 'effect', and thus what

it is to be a 'sufficient cause of an effect', depends upon one's theory of causation. Viewed from the perspective of one theory of causation, one might not see any 'causal gaps' in the physical domain that mental causes could need to fill. Viewed from the perspective of another theory of causation, the same will not automatically follow. Hence, I would suggest that the stronger one's formulation of the causal closure principle, and the broader one's understanding of causation, the harder it will be to provide a definition of the term 'physical' which is both non-trivial and plausibly causally closed.

How might the interactive mentalist respond to the causal closure argument? Well, of course, one might deny its causal closure principle. But, alternatively, rather than denying the causal closure principle, the interactive mentalist may seek to show that the premises of the causal closure argument are in fact consistent with interactive mentalism. One way of doing this is to deny the homogeneity of causation.

The premise of the homogeneity of causation is hidden within both strong causal closure arguments and the argument from causal overdetermination. With regard to the strongest causal closure argument, if all causation that is of kind x only involves physical causes, then it only follows that no effect has a non-physical cause, if all causation is of kind x. Similarly, if all causation that is of kind x only involves physical causes, then it only follows that no physical effect has a non-physical cause if the only causal difference that can be made within the physical domain is of kind x. With regard to the argument from causal-overdetermination, it may be true that every physical effect has a set of direct physical causes which together are causally-x sufficient for its occurrence, but it is only if all causation within the physical domain is of kind x that it follows that every physical effect has a set of direct physical causes which together are causally sufficient for its occurrence.

As we have seen, Crane (1995) raises the issue of the homogeneity of causation in relation to non-reductive physicalism, arguing that the non-reductive physicalist typically tries to address the problem of mental causation facing him, by denying the homogeneity of mental and physical causation. One of the aims of this discussion is to take Crane's observation and examine whether or not it can be used to cast light on the debate between the interactive mentalist and the psychophysical reductionist.

Here let me remove one potential point of misunderstanding. Let us assume a weak causal closure principle. One might argue that if the physical world is truly causally closed, then a physical effect will be causally overdetermined if a mental CP makes any kind of

causal contribution to it. If the mental cause does not causally overdetermine the physical effect then this means that there is a causal gap in the physical world — it is unimportant whether it is a gap in one kind of causation as opposed to another — the physical world is not truly causally closed. In other words, what the monist assumes when he advances the causal closure principle is that the physical world is causally closed in every plausible sense of the word 'cause' that is of relevance to the physical domain. Hence, the premise of the homogeneity of causation serves no purpose in the monist's argument. (This argument is similar to the one advanced by Kim in his criticism of Burge's (1993) discussion of explanatory exclusion (Kim 1999a, pp.64- 65)).

To reply, let us distinguish between x and y causation and thus deny the homogeneity of causation. Furthermore, let us say that both kinds of causation occur within the physical domain. The physical world may be causally x-closed, but given the above argument, the monist should respond that this does not mean that the physical world is causally closed. To be causally closed the physical domain must be both causally x and causally y closed. Now it is true that in order for a monistic conclusion to be entailed by the argument from causal overdetermination, the physical domain must be causally y closed as well as being causally x closed. However, in order to recognise that this is what the monist requires within his causal closure principle, attention must first be drawn to the possible heterogeneity of causation. Otherwise, from the fact that the physical world is causally x closed, the monist will automatically assume that the physical world is causally closed. That is, he will conflate causation with x-causation.

What it is for there to be a causal gap in a chain of causation, depends upon the analysis of causation that one is assuming. If mental and physical causation are heterogeneous, then from the aspect of physical causation there may well appear to be a seamless chain of physical causes. It is only once one recognises the heterogeneity of causation that one recognises that there are further causal roles to be played, and hence further kinds of causal gaps to be filled. And once one recognises the possible heterogeneity of causation, one needs to reassess what it is for the physical world to be causally closed. If there are additional causal roles for causes to play within the physical world, then the monist must show that physical causes can play these roles. Of course, one can amend one's causal closure principle so that the physical world is causally closed in every sense of the term 'cause' that is applicable to the physical domain. However, my point is that the more one liberates one's understanding of what it is to be a cause within the physical domain, the harder it is to provide justification for the causal closure principle.

To see what is involved in denying the homogeneity of causation, one must relate one's discussion to a specific theory of causation. Unless one considers what it means to say that something is a cause, one cannot adequately understand what it means to deny that everything is a cause in this sense. Indeed as we shall see, the importance that one attaches to the premise of the homogeneity of causation depends entirely upon one's analysis of causation. There is, for example, little motivation in the context of the mental causation debate in denying the homogeneity of causation, given a mere regularity theory of causation. This is because a regularity theory allows a very liberal understanding of what it is to be a 'cause'. The issue of the homogeneity of causation only becomes important when one's theory of causation offers a narrow interpretation of a 'cause', which is, furthermore, biased towards a physical model of causation. In particular, it shall be argued that if one assumes an energy transference theory of causation within the physical domain, this demands the denial of the homogeneity of causation if one is to avoid begging the question against the mentalist. This becomes particularly interesting once it is recognised that many of the monistic assumptions that further some kind of causal closure principle, make assumptions about psychophysical causation that are only plausible if one is assuming an energy transference theory of causation.

1.4 Outline

I begin by considering how to define the 'physical'. §2 and §3 are concerned with two different methods of defining the 'physical'. §2 considers the idea that there is a mark of the physical (where the physical is being understood in a broad sense). After rejecting this method of defining the physical, in §2.2, I go on to consider a priori arguments for causal closure principles, some of which rely on there being a mark of the physical. Various considerations lead me to conclude that these kinds of arguments are highly implausible and should be rejected. Any plausible argument for any type of causal closure principle must be based upon empirical considerations, more specifically empirical considerations from within physics. This leads on to the question of what 'physics' the monist is referring to, and the dilemma raised by Hempel, that there is in fact no physics to which he can plausibly refer. That is, that all definitions of the 'physical' lead to the conclusion that monism is either (probably) false, or must make physics too nebulous for useful discussion.

In §4, §5, and §6, I consider what kind of empirical support could be offered for the three strengths of causal closure principle. I hope to show that the monist's (implicit) acceptance of an energy transference theory of causation lies behind the most plausible arguments for all strengths of causal closure principle, and that although this theory of causation may offer a plausible account of purely physical causation, it should not be applied to psychophysical and purely non-physical causal interactions.

That an energy transference theory of causation provides the central support for causal closure principles is most obviously the case with strong causal closure principles. Given the rejection of an energy transference theory of causation, there is little, if any, reason to support a strong causal closure principle. However, the kind of appeals made by the monist to motivate the weak causal closure claim — the claim that mental causes will never be needed to account for physical effects — also implicitly appeals to assumptions about causation that are only plausible if one is assuming an energy transference theory of causation. I argue that it only seems plausible that the mental will never be needed to account for physical effects, because the energy transference theory of causation is being implicitly appealed to to determine what a causal gap in the physical domain must be like. Furthermore, it is being assumed that causal gaps in the determination of certain physical effects must be filled by mental causes in the same way that physical causes fill causal gaps, that mental causes must behave like forces, and that there must be a non-causal mechanism behind all such causal relations. If an energy transference theory of causation is rejected, these assumptions become implausible. I go on to argue that if one bases one's theory of psychophysical causation in a less physically biased theory of causation, then even the weakest kind of causal closure principle that would support monism loses much of its persuasive force.

2

A Priori Arguments for Causal Closure Principles

2.1 A Mark of the 'Physical'

One can distinguish between two different (although potentially inter-linked) methods of defining the 'physical'. Let us refer to the set of objects, properties and relations appealed to within the discipline of physics, as the set of 'physical a' objects, properties and relations. The set of objects, properties and relations appealed to within the higher-level sciences, shall be referred to as the set of 'physical b' objects, properties and relations. Lastly, the set of 'physical ab' objects, properties and relations consist in the conjunction of set a and set b.

Given that the substance dualist wishes to deny that all substances are composed of matter, when contrasting it with physicalism, it is standard practice to adopt a broad understanding of the term 'physical' in which the physical is identified with the physical ab. Hence, one method for demarcating the physical from the non-physical is to identify a (complex) set of exactly resembling CPs that all and only physical ab substances necessarily instantiate CPs from; the further question being whether mental substances fail to instantiate a CP from this set, and thus whether substance dualism is correct.

The second method is to identify a property-type as 'physical' if it is appealed to within the discipline of physics — it is after all physics which the physicalist considers to be authoritative. I consider this second method to be the most appropriate way of defining the 'physical', at least for the purpose of the mental causation debate. Here I wish to consider the main reasons why the first method fails.

In suggesting that the 'physical' refers to an 'extremely basic natural kind' Snowdon (1989, p.154), adopts the first method of defining the physical. According to him, the physical is a natural kind, because there is a potentially complex and arguably intrinsic property-type,

which all physical objects essentially share. Furthermore, this property-type marks off physical objects — it is in virtue of it, that there is a real difference between those objects that are physical and those that are non-physical. So as not to ignore traditional Aristotelian accounts of natural kinds that maintain that substances (in the sense discussed in Part One, §5.3) are special in having real essences, let us instead say that if there is a (complex) property-type that it is both necessary and sufficient that an object instantiates a CP from in order to be physical, then the term 'physical' refers to a natural category.

Let us for a moment assume that we have within our grasp the list of objects that a complete and accurate science refers to as 'physical ab'. One is faced with three possibilities. The first is that, as Snowdon wishes, there is a property-type that they share that is the 'essence' of being physical. The second is that there are a number of property-types that they essentially share, but none that is primary. The third is that in actual fact there is no property-type that physical objects essentially share, for there *need* not be a mark of the physical. Of course, if one subscribes to a semantical analysis of properties such as E4, it follows that there must be a set of exactly resembling CPs that all objects that fall under the predicate 'physical' instantiate a CP from. But as we have seen E4 is false. There need not be a (complex) type, corresponding to the predicate 'is physical' that all physical objects must instantiate a CP from. *s1* and *s2* may both be physical in virtue of a family resemblance.

Even if there is a mark of the physical, this is not to suggest that one can identify it or the method by which such an identification should take place. Just as Kant considered it a part of our concept of gold that it is a yellow metal, one may claim that it is part of our concept of a physical ab object that, for example, it is located in space. Guaranteed by the logic of the relevant concepts, it is an a priori truth that could not be discovered to be empirically false. If, however, the term 'physical' refers to a natural category, any a priori characterisation of it must be disregarded. Given a realist interpretation of natural kinds, they are groupings that exist independently of our conceptual and linguistic activities. If the physical is a natural category the same follows. Its identification will not depend upon laws of logic and definition, and therefore knowledge of whether or not an object is physical cannot simply be gained by ordinary use and mastery of the term. Indeed, it is precisely to disassociate themselves from the earlier materialists that did assume certain aspects of physical theory to be knowable a priori, that those supporting the ontological

superiority of the physical refer to themselves as 'physicalists' rather than 'materialists'.² The 17th century materialist placed a priori limits on what it was to be 'material', firmly anchoring themselves to traditional accounts of matter as inert, located and extended in space, and impenetrable. As physics has evolved, its understanding of matter has radically altered, leading to the subsequent rejection of 17th century materialism. To suggest that having property-type \underline{x} is an a priori requirement for being physical, just because it is currently considered to unite all physical objects, is to ignore the error in the materialist's method. If, as Snowdon's proposal suggests, it is the aim of physics to examine a certain level of nature searching for fundamental kinds defined by real essences, its understanding of the 'physical' should not be constrained by an a priori characterisation. One cannot rule out the possibility that empirical investigation may show that not all of the objects within this natural category instantiate a CP of this type, or indeed that one was mistaken in characterising the initial sample of objects as all instantiating a CP of this type in the first place.

Snowdon does not specify the theory of natural kinds that he adopts, however his suggestion that the physical is a natural kind is most plausibly interpreted along the lines of a Kripkean (1980)/ Putnamian (1975) account. According to them, not every a posteriori truth must be a contingent one. True identity statements constructed from distinct singular terms that are either proper names or indexicals are the clearest examples of necessary a posteriori truths. But identity statements involving natural kinds such as 'Water is H₂O' are also arguably necessary a posteriori truths. That water is H₂O could only have been an empirical discovery, however in all possible worlds in which water exists, water is H₂O. Kripke goes on to explain how natural kind terms are introduced into our language and how this can be in advance of any knowledge of the underlying structure of their referents. A community recognises a number of items that bear apparent similarities to each other. In virtue of these similarities, it is assumed that there is a natural kind that is common to them. Ostensively (or otherwise) the paradigm sample is identified and a term is introduced to refer to the kind. After the baptism, there remains the possibility that the sample does not share a common internal nature, and even if it does, at the time of reference fixing, the community may have no idea what it consists in. What the internal nature of the natural kind is, is an empirical question, which may be explored through scientific investigation. Until one has discovered the internal nature of a natural kind one's criteria for classifying items as belonging to it may be in error, as whether any or many of the property-types that figure in the concept of the natural kind actually characterise it, or

² For a discussion of this point, see Crane and Mellor (1990, pp. 186-7).

in Putnam's words, whether the stereotype of the term provides a good guide to its extension, depends upon the internal nature of the natural kind.

Can a similar account be given for the term 'physical'? For this to be the case, its extension must be fixed by ostension. The initial criteria for recognition of something as physical must be epistemically contingent. The original classification of objects as physical must be revisable in response to scientific investigation of this category. Finally, it must be possible to construct plausible twin earth cases for physical objects.

An account of the term 'physical' can be advanced that fulfils the first three of these criteria. Through our causal contact with various macro-objects such as tables and trees we recognise that these objects resemble one another in a number of important but very general ways. For example, they all possess a spatial location and are extended within space. In virtue of these similarities, it is assumed that they belong to the same natural category. The term 'physical' is introduced to refer to this category. Scientific investigation has led to the alteration of the stereotype with which its inquiries initially began. For example, current physics no longer considers all physical objects to possess a determinate location in space and it is highly probable that further adjustments will be made. However, that plausible twin earth cases can be constructed is questionable, for it is improbable that a (complex) property-type that can be non-arbitrarily singled out as essential for an object to be 'physical' has been identified, if indeed there is one.

If there is a mark of the physical, what might it be? Kripke and Putnam's suggestion that 'Water is H₂O' is a necessary truth, assumes that being a sample of the same substance as something consists in having the same chemical structure. Similarly, the claim that 'Gold has the atomic number 79' is a necessary truth, assumes that being a sample of the same chemical element as something consists in having the same atomic number. These essentialist claims are guided by the underlying principle that composition controls stereotype. Can similar essentialist principles be advanced with regard to the physical? Indeed, is it not the case that 'Water is H₂O' and 'Gold has the atomic number 79' are both just applications of the general claim that physical b items essentially have the microstructure that they have?

Let us assume that Kripke and Putnam are correct to maintain that composition controls stereotype, and that this is knowable a priori by philosophical analysis. Has the physicist yet discovered what the composition of all physical objects probably consists in? One can assert that physical objects larger than an atom are composed of atoms, and in a possible

world in which an object shared similar properties to those stereotypical of physical macro-objects but which was not composed of atoms, it would not be physical — an ethereal entelechy could at best have the appearance of a physical macro-object. However, not only is it improbable that we have yet discovered the ultimate constituents of the physical world, if indeed there are such things, but as entities at the bottom-most level will have no physically significant proper parts, rather than composition, one must identify some alternative essential feature. And once one begins to consider the kind of property-types invoked by physics, one realises that they are simply the wrong kind of thing to form natural categories. Theories of natural *kinds* require that the fundamental properties are intrinsic. But physics characterises its basic entities relationally, in terms of their causal relations with other objects, their propensity to interact with others. For example, as Bilodeau states, at the quantum level 'The particles are perhaps best thought of as units of causal influence' (1997, p. 225). Thus whether or not something is a muon depends upon the extrinsic relations it has with other basic particles. One does not learn from physics what the internal nature of this particle is, if indeed it possesses such a thing. Consequently, one should have serious doubts about whether the 'physical' could be a natural category, as the objects that physics describes appear to be of the wrong kind.

Hence our rejection of a semantical criterion allows us to recognise that the physical need not be a natural category, and consideration of physics should give us reason to question that it could be. Finally, even if one could identify a mark of the physical, although this will allow one to distinguish a substance monism from a substance dualism, it will not allow one to distinguish a property pluralism from a property monism. To do this one must be able to demarcate physical properties from non-physical properties. Now Snowdon (1989, p. 154) indicates that by demarcating physical objects from non-physical objects one thereby demarcates physical properties from non-physical properties, as he considers a property to be physical if it can be instantiated in a domain consisting only of physical objects. Certainly, this would be true if, for a property to be physical, was simply for it to be a way that a physical object is. But this is clearly false. Properties are not automatically physical because they are ways that physical objects are.³ Mental properties may be properties of the brain, but to claim that mental properties are therefore 'physical' because they are ways that physical objects are, is to trivialise one's understanding of a physical property, and thus the debate between pluralism and monism.

³ Note furthermore, that Lowe (1989, ch. 2) considers persons to possess physical properties. Consequently, not everything that has physical properties is part of this single macro-micro hierarchy of objects.

This issue must be distinguished from that of whether by determining what it is to be a physical object, one thereby determines the different kinds of property that a physical object will instantiate. Essences are often considered to determine (some of) the properties of objects belonging to the kinds of which they are essences. Therefore if the essence of being a physical object is having a certain microstructure, plausibly having such a microstructure nomologically determines a range of other property-types, which are thereby derivatively essential to physical objects. Thus it is arguable that in a possible world nomologically identical to our own, objects possessing the relevant microstructure must have the properties of mass, position, and momentum at the micro-level, and at the macro-level the properties of extendedness, impenetrability and inertness. However, once again, one would be wrong to assume that all properties that are derivatively essential to physical objects are therefore physical. From the fact that all substances with a certain microstructure are essentially conscious, it does not follow that consciousness is therefore a physical property.

For these reasons, I would suggest that even if there is such a thing as the mark of the physical, identifying it will not enable us to identify what it is for a property to be physical, which is precisely what is required to distinguish monism from pluralism. To pick out the properties that are physical a, one must instead look to those appealed to by physics. But as attempts to limit what physics can be by fixing its subject matter are doomed to failure, this raises the question of how one should understand physics.

2.2 A Rejection of A Priori Arguments for Causal Closure Principles

Before we move on to consider this more plausible method of defining the 'physical', it will first be convenient to rule out one kind of argument often invoked to advance strong causal closure principles. These kind of arguments appeal to a priori reasons to rule out the possibility of either non-physical causation or psychophysical causation. I discuss them here, firstly, because they are commonly based around the assumption that there is some mark of the physical (and hence assume the above method of defining the physical). Secondly, having dismissed any a priori arguments for the strong causal closure arguments, this leads one to recognise that, as with the weak causal closure principle, any plausible arguments for strong causal closure principles will be based upon empirical considerations, more specifically empirical considerations from within physics. Hence, the

question of how one should understand physics is crucial, not only to the argument from causal overdetermination, but also strong causal closure arguments.

There is no plausible a priori reason to support a causal closure principle. Contrary to this, one may attempt to defend the claim that 'No effect has a non-physical cause' by appealing to the fact that some theories of causation, for example, Hume's theory of causation, incorporate the requirement that cause and effect must be spatially contiguous. (See Part One, §1.2). If spatial contiguity is a necessary feature of the causal relation, and furthermore, if it was the mark of physical substances that they possessed a location in space, then this would rule out the possibility of non-physical substances being involved in causal relations, for clearly in order for one substance to be spatially contiguous with another it must first be located in space. Therefore those substance dualists who combine the claim that mental substances lack a location in space with the claim that the properties of mental substance make a causal difference are in error. Hume, for example, faced this dilemma (1738, Book 1, Part IV, S5) and it was for this reason that the requirement of spatial contiguity was later omitted in his *Inquiry* (1748).

But this argument fails at every stage. Firstly, as discussed, it is debatable that there is a mark of the physical, and certainly that it is that of having a spatial location. The monist cannot set a priori limits on what it is to be 'physical' and in the light of developments within physics the claim that all physical objects necessarily possess a determinate location in space is particularly contentious. Equally questionable is whether there can be action at a distance within the physical domain, and hence whether physical causation must always be local. Hence, if spatio-temporal contiguity is a necessary feature of causation then, in actual fact, this presents a potential difficulty not only for the mentalist but also the monist. On the other hand, it is questionable that spatial locatedness is a sufficient property for being a physical substance; Some philosophers understand minds to be essentially similar to Cartesian souls, except that they are spatially located. (For example, see Quinn (1997)). These entities should not be identified as physical, merely because they are spatially located.

Rather than bringing into question the suitability of certain mental and physical substances as causal entities, one should instead question any analysis of causation that assumes that it is an a priori truth that cause and effect must exhibit spatio-temporal contiguity. Although it may be an a posteriori necessary truth that all causal relations in this world do exhibit spatio-temporal contiguity, it is not part of the concept of a causal relation that cause and effect are spatially contiguous. Hence, whether or not it is the case that all

causes and effects are in fact spatially-temporally contiguous, is to be settled by empirical considerations. (For a discussion of this point, see Part One, §1.2).

There is of course, the further question of how mental states could be causes of spatially located entities if they lacked a location in space — obviously not by proximate causes and neither by action at a distance. McGinn (1997) believes this makes mental causation mysterious — ‘Recent philosophy has become accustomed to the idea of mental causation, but this is actually much more mysterious than is generally appreciated, once the non-spatial character of consciousness is acknowledged. To put it differently, we understand mental causation *only* if we deny the intuition of non-spatiality’ (McGinn (1997, p. 100)). But McGinn neglects to consider that *how* one state causes another is arguably a notion that appears to escape us, even when the states concerned are both spatially located. The search for an understanding of how one property-instantiation causes another is according to many theories of causation (the regularity theory, the counterfactual theory, neo-Humean theories etc.), a misguided one, that should not be raised. And those theories of causation that do attempt to offer an explanation of how a cause produces its effect, certainly do not require that the causal relata are spatially located (with the possible exception of the energy transference theory).

Finally, even if it was the mark of physical substances that they were spatially located, and it was an a posteriori truth that cause and effect must be spatially contiguous, this would still not entail that ‘No effect has a non-physical cause’, for it would only rule out an interactive substance dualism (that maintained mental substances lacked a spatial location), not an interactive property mentalism. Causation is a relation between property-instantiations. Hence, what is of concern is whether only physical property-instantiations can have effects. And, as we have seen, just because a property is instantiated by a physical substance, it does not thereby qualify as physical. Even if all substances are physical (that is, according to this argument, spatially located), there may be non-physical property-instantiations and hence non-physical causes. To suggest otherwise would be to reduce interactive property mentalism to physicalism. Hence, even if it is the mark of physical objects that they are located in space, and only those objects that are located in space can enter into causal relations, it does not follow that there are no non-physical causes.

The same kinds of objection can be raised against any attempt to motivate some a priori reason to suggest that non-physical causes could not be causes of physical effects. Hence, take the commonly raised objection that Cartesian dualism makes psychophysical

causation unintelligible because mental and physical substances instantiate wholly opposing properties, and entities that are of a different kind cannot enter into causal relationships with one another. To this one must ask why one should assume that only things of the same, or similar, kinds can enter into causal relationships with one another? Knowledge of the properties belonging to two substances does not provide any a priori reason to suggest that they could not enter into a causal relationship with one another. At best the claim that only like causes like, can be a consequence of some empirical theory.⁴ Hence, if there is any support for the claims that 'No effect has a non-physical cause' or that 'No physical effect has a non-physical cause' it must be based upon empirical rather than conceptual considerations.

⁴ This is essentially Hume's point that we cannot determine a priori what causes what. (1738, Book 1, Part IV, Section V). For further discussion of this point see Lowe (2000b, pp. 21-24) and Broad (1929, pp. 97-8).

3

An Appeal to the Discipline of Physics

3.1 Sturgeon's Dilemma

The physicalist requires that the 'physical' domain is causally closed. But equally, to advance a non-trivial physicalism, one's definition of the term 'physical' must adequately demarcate the physical from whichever category one is contrasting the physical with. Hence, to provide a non-trivial monism, physical a properties must not include paradigm examples of physical b properties, and in order to provide a non-trivial anti-mentalism, physical ab properties must not include paradigm examples of mental properties. The question is, could a physical domain exist that is both causally closed in the required sense and non-trivial? Obviously, the stronger one's causal closure principle, the less plausible this will become. For the purpose of defining the 'physical', I shall assume the weak kind of closure principle that is appealed to within the argument from causal overdetermination, although it should be clear that similar arguments regarding whether to identify 'physics' with current or future physics apply to any causal closure principle that appeals to current physics in order to motivate its claims.

Now there is an immediate problem with the argument from causal overdetermination's understanding of the 'physical'. The plausibility of its causal closure principle and the correctness of its assertion that mental causes have physical effects, both depend upon how the term 'physical' is understood. But Sturgeon (1998) argues that these two premises require different understandings of the term 'physical' in order to be plausible, hence leading one to doubt the overall plausibility of the argument from causal overdetermination.

Its closure principle requires that every physical effect has a direct, sufficient physical cause, hence entailing that one never needs to appeal to non-physical causes in order to provide a direct sufficient cause for every physical effect. But this is implausible if the 'physical' includes physical b categories — biology needs to appeal to chemistry, and chemistry needs to appeal to physics, in order to provide a complete causal account of

their effects. It is physics that the monist considers to be autonomous. Consequently, its premise of causal closure requires the identification of the 'physical' with the physical a.

But the argument from causal overdetermination is also committed to the existence of psychophysical causation. Its claim is not that mental causes have physical a effects. The physical domain in which mental causes are commonly thought to have effects is the neurophysiological. Therefore, the premise that there is psychophysical causation assumes a broad understanding of the physical which includes the physical b. So to be plausible, the argument from causal overdetermination has to appeal to two different understandings of the term 'physical'.

Now the dilemma that Sturgeon raises does have an obvious solution. What the physicalist needs is an understanding of the relationship between the physical a domain and the physical b domain, which would allow him to assume a causal closure principle according to which: 'Every physical ab effect has a set of direct physical a causes which together are sufficient for its occurrence.' That is, a relationship between the physical a domain and the physical b domain, which would allow that the physical ab domain as a whole is causally closed in virtue of the fact that the physical a domain is causally closed. As it is the physical b domain in which mental states are commonly thought to make a causal difference, Sturgeon's dilemma would then be resolved.

What sort of relationship must there be between the physical a domain and the physical b domain for this to be the case? Well, given that it is the physical ab domain in which physical b causes are thought to have effects, if every physical ab effect had a direct sufficient physical a cause, this would imply that any causal difference that physical b properties make is nothing over and above the causal difference that physical a properties make.⁵ Given our discussion of property analysis, and the acceptance of a causal analysis of CPs, I would suggest that if physical b causes are nothing over and above physical a causes, this must mean that physical b properties are nothing over and above physical a properties. That is, the set of physical ab properties is identical with the set of physical a

⁵ The suggestion that physical b properties are causally nothing over and above physical a properties could be established by applying the argument from causal overdetermination to the descending levels of properties within the physical b domain. Hence, if, via the argument from causal overdetermination, one could establish that chemical causes are nothing over and above physical a causes, and that biological causes are nothing over and above chemical causes, one would then have shown that biological causes are nothing over and above physical a causes. As Papineau suggests, 'The causal argument provides a schema that delivers physicalism for other special subjects as readily as for the mental' (2000, p. 204, fn. 5).

properties. So one way of solving Sturgeon's dilemma is to maintain that the properties appealed to within chemistry, biology, etc. reduce to properties appealed to within physics.

Note that similar arguments to the one that Sturgeon raises, also arise with regard to the stronger causal closure arguments. Hence, take the second strongest causal closure argument. Any motivation gained for the causal closure claim that 'No physical effect has a non-physical cause' must, as we shall see, be based upon an appeal to physics. Hence, its causal closure principle must identify the physical with the physical a. On the other hand, it also appeals to the premise of psychophysical causation. Hence, this causal closure argument is appealing to two different understandings of the term 'physical'. Once again, one solution is to identify the physical ab domain with the physical a domain, from which it would follow that 'No physical ab effect has a non-physical a cause.'

Now this general kind of response to Sturgeon's dilemma requires one to assume a causal monism; the only properties that make a causal difference (within the physical ab domain) are physical a properties. Therefore, if there were emergent chemical or biological properties, this would give us a general reason to question the causal closure arguments that anti-mentalists advance. But not everyone will be happy with this response to Sturgeon's dilemma. Some may want to reject mentalism by way of a causal closure argument, whilst maintaining that there are emergent physical b properties. If the above is the only way of responding to Sturgeon's dilemma, this possibility is ruled out.

However, there is an alternative response to Sturgeon's dilemma that caters for this possibility. This is to reject Sturgeon's narrow interpretation of the physical within the causal closure principle. Here, I have in mind the definition of the 'physical' advanced by Papineau in 'The Rise of Physicalism'. Papineau (2000, pp.183-4) identifies the 'physical' as a contrastive term. This offers an alternative approach to Sturgeon's dilemma because the same understanding of the 'physical' can be appealed to within both the causal closure principle and the premise of psychophysical causation. Hence, let us consider the argument from causal overdetermination (although this way of defining the physical has general application within all causal closure arguments). If, for the purpose of the mental causation debate, one identifies the 'physical' with the non-mental, that is, one identifies 'physical' properties with properties that can be identified without using specific mental terminology, then the argument from causal overdetermination is as follows;

1. Every non-mental effect has a set of direct non-mental causes, which together are sufficient for its occurrence.

2. Mental causes have non-mental effects.
3. There is no systematic causal overdetermination.

Mental causes (that have non-mental effects) must be identical with non-mental causes.

It is important to note that this should not be seen as suggesting that the first premise requires a narrow understanding of the non-mental (the non-mental that is physical a), and the second premise requires a broad understanding of the non-mental (the non-mental that includes the physical b). If this were the case, then, this would not offer an alternative solution to Sturgeon's Dilemma.

Rather, both premises allow a broad interpretation of the non-mental. This is because, even if individual disciplines, such as the biological and the chemical, that fall into the category of being 'non-mental' are not autonomous, the anti-mentalists could argue that it will remain true that they will not need to appeal to mental properties in order to be causally complete. This way of formulating the causal closure principle is specifically targeted upon the question of whether or not there are emergent mental properties. It is silent about whether there are non-mental emergent properties. I shall return to this method of defining the 'physical' in my discussion of Papineau in §3.6. For the moment, I shall adopt the first kind of solution to Sturgeon's dilemma, and hence shall assume a causal closure principle which maintains that; 'Every physical ab effect has a set of direct physical a causes which together are sufficient for its occurrence.' However, my overall conclusion will be that the contrastive method of defining the 'physical' provides a more plausible means of advancing a causal closure principle against mentalism.

3.2 Hempel's Dilemma

One can distinguish between three different general understandings of physics that are appealed to by the monist. The first identifies physics with current physics.⁶ The second identifies physics with some future 'completed' physics. The third offers a definition of physics which renders the causal closure principle trivially true; physics is causally closed by definition.

⁶ Of course, the monist is not concerned with every part of current physics, but only those fundamental parts that examine the micro-constituents of macro-systems, namely quantum mechanics and particle physics.

Before we can discuss these various understanding of physics, it is first important to explain what I mean by 'completeness'. Like Melynck (1997), I understand physics to be 'complete' if and only if it is ontologically complete. Hence, for current physics to be ontologically complete, its list of physical a phenomena must not require any additions or modifications in the light of future discoveries. If, for example, an undiscovered property exists that current physics would classify as physical a then current physics is ontologically incomplete⁷.

This understanding of completeness is purposefully consistent with the possibility of a physics that is complete but which is not able to provide a complete causal account of every physical b effect, nor for that matter a complete causal account of every physical a effect. Why do I distinguish the question of whether physics is complete, from the question of whether physics is autonomous? Well, it is because I do not think that one should assume that just because a property is invoked by physics in order to causally account for a physical a effect, this therefore entails that it is physical a. To justify the ontological superiority which the monist gives to the physical, his understanding of the 'physical' must reflect the physicists', as it is physics that is taken to have some more or less privileged claim to truth. But, if, for example, it was discovered that consciousness caused the collapse of the wave function, then although the collapse of the wave function is a physical a state, I do not think that one can therefore assume that the physicist would classify consciousness, the cause of this physical a effect, as physical a. It is certainly not obvious that the physicist should or would consequently set about investigating consciousness. This would seem to be the role of psychology, not physics.

For similar reasons, I like Papineau's (1990, 1993) trivial definition of the 'physical' (to which I shall return in §3.6) even less. According to Papineau, the function of physics is to study the causes of paradigm physical effects. But the *physicist* doesn't represent himself as studying *everything* that has causal effects within the domain of paradigm physical objects. If a psychological or economical property was found to cause a paradigm physical effect, it is unlikely that it would be the *physicist* who would set about its study. Papineau has, I think, confused the aim of the physicist with that of the monist. The monist's primary goal is to establish that every paradigm physical effect has a direct sufficient physical cause, but the aim of the physicist is not therefore to study the set of properties that meet this requirement.

⁷ See Melynck (1997, p.623, fn.7.). Melynck's definition of 'physics' shall be considered in §3.3.

With this weaker understanding of completeness in mind, I shall now outline Hempel's dilemma. Let us first consider those monists that appeal to current physics in order to define their position. Let us refer to monists as c-monists if they accept C1, where C1 is the claim that:

C1: A property is physical a, if and only if it is essentially referred to within the theories and laws of current physics.

Current physics, the c-monist will argue, never needs to look elsewhere in order to causally account for physical a effects. Given c-monism, it is therefore plausible that every physical a effect has a set of direct physical a causes which together are sufficient for its occurrence. However, in the first place, given the first solution to Sturgeon's dilemma, the c-monist does not merely require that every physical a effect can be accounted for in terms of physical a causes. In addition, current physics must also be able to provide a complete causal account of all physical b effects. There is little, if any reason, to think that current physics is causally closed in this wider sense.

Furthermore, it is highly improbable that physics is complete. An examination of the history of science yields countless rejections and modifications of scientific theories that were at some time considered to be true or partially true. Therefore, if the 'current' theories of past sciences are representative of today's physics, one should conclude that current physics is most probably not an accurate representation of the facts. Pessimistic induction provides one with good grounds for believing that current physics is unlikely to be complete or wholly accurate. The sceptical doubt that the argument from pessimistic induction raises regarding the completeness of current physics, generates a corresponding scepticism with regard to any monism that frames the distinction between the physical a and the non-physical a in terms of it, either for the purpose of defining monism or for defending the causal closure principle.

But if the monist cannot identify physics with current physics, then to which physics can he refer? A dilemma, proposed by Hempel (1980, pp. 194-5), suggests that there is no physics to which the monist can plausibly turn. All definitions of the 'physical' lead to the conclusion that monism is either (probably) false or has a truth-value that is indeterminable by us. As current physics is most probably incomplete, a monism framed in terms of it, is most probably false. Given that one cannot appeal to current physics, a plausible (non-trivial) monism must therefore be based upon a physics that does not yet

exist. Let us call those monists who identify the physical a with the contents of some future physics 'f-monists'. F-monists accept C2, where C2 is the proposition that:

C2: A property is a physical a property, if and only if it is essentially referred to within the theories and laws of a future physics that is complete.

Future physics is by definition complete. But the problem for the f-monist is that as one does not have epistemic access to the physics to which it appeals, one is unable to determine the content of its laws. Consequently, one cannot assess whether such a physics needs to appeal to non-physical a causes in order to account for physical a effects. Neither can one assess whether every physical b effect could be causally accounted for by physical a states. Therefore, one cannot hope to assess the plausibility of a monism framed in terms of it.

Are there any plausible arguments to suggest that either the c-monist or f-monist can respond to Hempel's dilemma, thus providing one with an understanding of the term 'physical' suitable for advancing a non-trivial monism? In §3.3 and §3.4, I consider Melynck and Smart's different methods of defending c-monism, and suggest that both are implausible. In §3.5 I consider how one might defend an f-monism. Then in §3.6, I consider Papineau's suggestion that one can lay down a definition of the physical a from which the satisfaction of causal closure trivially follows, and then by adding a further empirical exclusion claim, advance a plausible understanding of monism. I do not think that Papineau's trivialisation of the causal closure principle serves any purpose. The important fact that Papineau's discussion of the physical does show, is that one can focus one's attention on the sub-section of non-physical properties that is of relevance to the mental causation debate. This consideration can be detached from Papineau's trivialisation of the term 'physical', and incorporated into an f-monism, to yield a plausible understanding of the 'physical' suitable for the monist.

3.3 Melynck's Appeal to Current Physics

Melynck (1997) defends a c-monism. C-monism is held in question, Melynck argues, because of the argument from pessimistic induction (1997, p. 623), but in fact, this argument need not lead to the abandonment of c-monism. Pessimistic induction does not entail that current physics is incomplete (where completeness is to be understood in the

sense that I am assuming), only that it probably is. It therefore only raises a problem for the c-monist, if it is necessary for him to assign a high probability to his position, and Melynck argues that this is not the case. According to Melynck, to be a c-monist it is only necessary that one believes c-monism to be more probable than any of its relevant rival theories. This is consistent with assigning c-monism with a low probability.⁸

Given Melynck's proposal, the c-monist must maintain that more probable than any relevant rival hypothesis is the hypothesis that current physics is complete. The question is how can the c-monist justify assigning a greater probability to this hypothesis than any of its relevant rivals? Melynck provides a number of conditions that a 'relevant rival hypothesis' must fulfil:

'Hypothesis H1 is a relevant rival to H2 if and only if (a) H1 is sensibly intended to achieve a significant number of H2's theoretical goals; (b) the hypotheses, H1 and H2, fail to supervene on one another; and (c) H1 has actually been formulated.' (p. 626)

With these restrictions, Melynck can rule out the most obvious counter-example to his argument. It logically follows from Melynck's argument that a monist assigning a low probability to the hypothesis that current physics is complete, will assign a higher probability to its negation. However, the claim that 'Current physics is not complete' would not be allowed as a relevant rival hypothesis as it does not fulfil criterion (a). Without supplementation simply denying that current physics is complete cannot help one to identify the physics that actually is complete.

But even with these restrictions in place, Melynck's argument is fatally flawed. His mistake is to assume that reasons for questioning the completeness of current physics arise only from considerations of pessimistic induction, when in fact an examination of current physics *itself* leads to the conclusion that it is incomplete. For example, a good reason for thinking that current physics is incomplete is that the guiding principles that underlie the unification of quantum field theory and the theory of General Relativity are unknown. Physicists neither assume that current physics is on the brink of this discovery, nor that upon such a discovery a major upheaval within physics will not occur. As Weinberg suggests, providing a unified theory of forces will 'probably not be possible without

⁸ This is to take what Melynck refers to as a 'Scientific Realist attitude' towards monism. (1997, p.625). Whether this view is in fact represented by scientific realists will not be directly brought into question, as regardless of one's verdict it is implausible within monism.

radically new ideas.' (1999, p. 74). It is highly probable that in order to achieve such a unification, a new elementary category of entity, often referred to as Higgs particles, will have to be postulated. Furthermore, if proponents of the superstring theory are correct in thinking that they have found an acceptable bridge between gravity and quantum mechanics, this will entail a radical revision of physics' current ontology. It is therefore reasonable to conclude that current physics is incomplete. And for this reason, a monism that incorporates the claim that current physics is complete should be abandoned.

One may respond that this is to ignore Melynck's central argument; the monist can consistently assign a low probability to the hypothesis that current physics is complete so long as it is higher than that of its relevant rival hypothesis. However, this is to miss the point. Whilst the argument from pessimistic induction only suggested that it was highly improbable that current physics is complete, current evidence suggests that it is plainly false to think that it is complete. For this reason, I would suggest that Melynck's defence of current physics should be rejected. It also should be noted that Melynck does not even begin to tackle the question of how one might justify the claim that current physics can provide a complete causal account of every physical b effect as more probable than any of its relevant rival hypothesis. Melynck's defence of c-monism is therefore also incomplete. Finally, one may question Melynck's method of defining the term 'physical'. In defining the term 'physical' in the way that he does, Melynck's assumption is that monism should be assessed in the same kind of way that scientific hypotheses are. However, monism is not a scientific hypothesis, but a philosophical doctrine. Consequently, it is far from evident why monism should be assessed in this way.

3.4 Smart's Appeal to Current Physics

From the fact that all monists appeal to the causal closure of current physics in order to justify their position, it would not be presumptuous to assume that such a position follows from a detailed survey of quantum physics, the presently accepted description of the physical domain. But surprisingly, examinations of quantum mechanics are rare within the philosophy of mind.⁹ Although, for example, the argument from causal overdetermination, and therefore the ontological primacy of physics, plays a central role within Papineau's 'Philosophical Naturalism', what quantum physicists themselves have to say about the matter is all but ignored. Similarly, in the 'Philosophy of Mind' Smith and Jones maintain

⁹ There are notable exceptions e.g. Hodgson (1991) and Stapp (1997).

that 'contemporary science is thoroughly wedded' (1986, p. 59) to micro-reduction and the causal closure of physics, and that if the dualist rejects these claims then he 'cannot really complain if the scientist laughs his armchair speculations out of court.' (1986, p. 59). Yet Smith and Jones omit any critical discussion of whether the quantum physicist rejects such claims as ludicrous. Equally, discussions of c-monism focus wholly upon the issue of whether or not current physics is complete, neglecting to discuss the critical question of whether or not quantum, and therefore current, physics is committed to any form of causal closure principle that would support monism.

Unfortunately, reasons for this omission are rarely given by the monist himself. I consider there to be four likely explanations. The first is that the monist believes that within quantum physics, causal closure principles (at least of the weaker variety) serve as an unquestionable premise and therefore detailed discussion of quantum physics is unnecessary. Alternatively, although not serving as a premise, the monist may consider all physicists' theories to result in the causal closure of physics. The third explanation is that as there is no general consensus among physicists regarding its interpretation, the monist can ignore quantum mechanics. Thus Putnam comments 'I shall follow the materialist in ignoring quantum mechanics since it has no generally acceptable interpretation of the kind the realist advocates...' (1982, p. 230). But, of course, such a stance is only acceptable if the monist can assume that the physicists' eventual verdict will be compatible with that of his own. The fourth explanation is that the monist ignores quantum physics because he does not consider it to be the physics that is of relevance to his discussion. To return to the problem of the unification of physical forces, it appears irrelevant to the mental causation debate whether or not physics will need to appeal to Higgs particles. Even though it would require the rejection of the completeness of current physics, it should not lead one to reject the central principles underlying monism. Firstly, these particles are decidedly physical, and secondly, the monist would argue that they are causally isolated from any domain that is of causal relevance to the mental. The consideration that this is true of quantum physics as a whole, provides new hope for c-monism. Unquestionably, new entities and properties will be appealed to at the micro-level of reality but the incompleteness of quantum mechanics will not automatically entail the incompleteness of the current physics that is of relevance to the mental causation debate.

Smart (1978) adopts the last kind of position, when defining the 'physical'. Like Melynck, Smart maintains a c-monism, but rather than considering the completeness of physics as a whole, he focuses upon the particular subset of physical properties which he considers to be of relevance to the mental causation debate. Distinguishing between the 'physics of

bulk matter' and the 'physics of elementary particles', he argues that whilst the latter may not be complete, the former (probably) already is. Certainly, the physicist does not have complete knowledge of elementary particles and physics will need to undergo many fundamental changes before this is gained, but these alterations will not disturb the physicist's understanding of bulk matter. And Smart suggests that it is the 'physics of bulk matter' which is of importance to the philosophy of mind. If mental properties are causally related to physical properties then they are causally related to properties belonging to macro-objects at the neurological level, or in other words properties of 'ordinary matter' (1978, p. 341). C-monism is therefore only tied to a part of current physics, and it is a part that is (probably) complete. If Smart is correct, then the question of concern for the mental causation debate is whether the physics of bulk matter allows the plausible advancement of the causal closure principle. Smart would obviously suggest that it does.

But Smart's definition of the 'physical' cannot plausibly be appealed to by the monist unless two further premises are assumed. Firstly, one must assume that the c-monist can, and indeed should, prise apart the physics of bulk matter from the physics of elementary particles. And secondly, that if such a division can be made, it is only the physics of bulk matter that is of relevance to the mental causation debate. The first premise may be considered to be supported by the Copenhagen interpretation of quantum mechanics. This considers measuring systems to be classical systems that operate according to classical physics, and micro-systems to be quantum systems that operate according to quantum theory. Consequently, a sharp divide is assumed between those systems to which quantum physics applies (e.g. atoms and their constituents) and those to which classical physics applies, which may be thought to correspond to the division made by Smart between the physics of elementary particles and that of bulk matter. But for it to support Smart's position, the Copenhagen division between classical and quantum systems must correspond to a division between macro and micro systems, which is debatable.¹⁰ Furthermore, that separate dynamical theories are to be applied within the macro and micro domains faces serious difficulties in its articulation. If there is a division between the field in which quantum theory (or the physics of elementary particles) operates and the field in which it does not, where is it to be drawn? One can provide clear examples of macro-objects such as doors and tables, and of micro entities such as electrons and quarks, but any dividing line is difficult to determine without becoming arbitrary.

¹⁰ For example, see Bilodeau (1997, p. 226).

Furthermore, whether the division that Smart makes is one that a monist would wish to support, and strictly speaking whether he would be consistent in doing so, depends upon the relationship between the physics of elementary particles and the physics of bulk matter. To support his position, the monist has always been quick to appeal to the enormous success of understanding macro-systems in terms of the micro-systems from which they are composed. In rejecting the universality of quantum physics, Smart effectively abandons this physicalist program. The micro-objects that compose macro-objects operate according to quantum mechanics, but macro-objects do not. Furthermore, if there are two distinct dynamical theories in operation in different physical domains of the world, this logically entails that there must be at least two fundamentally different kinds of physical phenomena within the world. If one maintains that there is more than one domain of independently causally efficacious properties then one is rejecting monism. The temptation is, of course, to resolve both of these potential difficulties by suggesting that the physics of bulk matter and the physics of elementary particles are intimately connected, and that the causal efficacy of bulk matter is nothing over and above that of elementary matter. But clearly, Smart cannot be allowed such a move. Smart maintains that all that matters to the mental causation debate is the physics of bulk matter. One does not have to wait and see which categories a physics of elementary particles will contain. But whether or not the causal efficacy of bulk matter is indeed nothing over and above that of elementary physics, depends precisely upon the kind of properties that the latter includes.

Finally, even if one was to accept the distinction between the physics of elementary particles and the physics of bulk matter, it is questionable whether all that matters to the mental causation debate is the physics of bulk matter. Quantum mechanics is arguably of great significance to the discussion of emergent properties in general, and of emergent mental properties in particular. Firstly, it is arguable that the quantum world is not a hierarchically structured one. Regarding Bell's inequality, the Aspect experiment shows that contrary to the EPR conclusion performing a measurement of the position or momentum of one particle enables one to predict with certainty the result of carrying out a measurement of position or momentum of another, regardless of how far apart they might be. This suggests that at least when *micro-states* are measured their behaviour is a function of the behaviour of the whole system, rather than that of its component parts. This necessitates a radical restructuring of our fundamental ideas about the nature of physical reality and the relation of part to whole, a restructuring that cannot be confined to the atomic scale. Naturally, this is of significance to the general question of whether there

are emergent properties. Secondly, Wigner (1961), and London and Bauer (1983), consider consciousness to cause the collapse of the wave function, and hence would dismiss Smart's assumption that if the mental makes a causal difference within the physical domain, then the physical domain is automatically to be identified with the physics of bulk matter.¹¹ Hence, Smart's method of defining the 'physical' should be rejected.

3.5 Future Physics

Whilst c-monism must be rejected because the physics to which it refers is incomplete, f-monism faces the other horn of Hempel's dilemma. How can one hope to justify a monism based upon a physics which does not yet exist? In particular, how can one know that future physics will not need to appeal to mental causes in order to causally account for physical ab effects, and hence how can one justify the kind of closure principle appealed to within the argument from causal overdetermination? The f-monist makes physics too nebulous for useful discussion.

Whatever the method that the f-monist appeals to, it will be based upon the assumption that current physics does not provide one with any reason to think that one will need to appeal to non-physical a categories in order to explain physical ab effects. For obviously, it is only then that it is hard to conceive how an extension of physics, possessing a similar structure, could require one to appeal to non-physical a categories in order to explain physical ab effects. Given this assumption, one could then argue that current physics provides a reliable guide as to what complete physics will be like; that complete physics will bear important similarities to current physics not only in its research methods, experimental techniques and its standards for evaluating the work of other scientists, but also the kinds of properties contained within it. And, therefore, if on the basis of current physics, the causal closure principle is plausible, this provides us with a reason to think that the same will be true of complete physics. More specifically, if on the basis of current physics, one does not need to appeal to a kind of property in order to causally account for physical ab effects, this provides us with a good reason to think that a completed physics will not either.

¹¹ I shall return to this proposal in § 6.5.

But obviously, whether this is plausible depends entirely upon the relationship between current and complete physics. Clearly, it would be to beg the question if the f-monist was to assume that however radical the shift in the concepts, theories and laws in physics, and therefore its entities, properties and relations, the one fact that would not alter is that properties such as the mental and economical, the chemical and biological would not *need* to be appealed to by physics in order to explain physical a effects. Of course, if the structure that complete physics invokes is continuous with, albeit an extension of, the structure that current physics invokes, then current physics would provide a reliable guide to what complete physics would be like, and hence one would have good grounds for suggesting that complete physics will be causally closed in any sense that current physics is. However, it is precisely what the relation will be between current and complete physics that the f-monist is unable to gain epistemic access to.¹²

3.6 Papineau and 'Closed Physics'

Rather than identifying the physical a with a current but incomplete physics, or a future but potentially causally open physics, one may lay down a definition of the physical a from which the satisfaction of causal closure trivially follows. Let us assume the kind of causal closure principle appropriate to the argument from causal overdetermination, primarily because Papineau (1990, 1993) has done the most to defend this method of defining the physical, and it is this kind of causal closure principle that he assumes.¹³

According to Papineau, a property is a physical a property only if it is required in order to provide a (direct) sufficient cause for a *paradigmatic* physical effect. Physics simply means 'the science of whatever properties are needed for a complete set of laws covering such effects as stones falling, darts hitting boards, etc' (1990, p. 70). Let us call the set of properties required in order to explain paradigmatic physical effects, the set of properties belonging to 'closed physics' and those who define the physical in this way as t-monists. T-monism has close links with f-monism, as Papineau obviously requires the 'complete

¹² And even if current and complete physics are not wholly discontinuous, one might ask why one should assume that complete physics is causally closed? Quantum mechanics' holistic view of the world, removes any assurance the f-monist may have had in the claim that a future physics will not *need* to appeal to non-physical causes.

¹³ Papineau maintains a non-reductive physicalism rather than a psychophysical reductionism.

set of laws' to be the correct set of laws. However, it does not encounter the problem facing f-monism, as from its completeness, causal closure trivially follows.

Does this understanding of the physical allow one to adequately formulate monism? Given that one understands the causal closure principle as the claim that; 'Every *paradigmatic* physical effect has a set of direct physical a causes which together are sufficient for its occurrence', then Papineau would be the first to admit that t-monism trivialises the causal closure principle — indeed this is the whole point! But in trivialising the causal closure principle, Papineau has been criticised of trivialising monism.¹⁴ Non-trivial monism relies on there being a contrast between the mental and the physical a. But if, as the interactive mentalist believes, mental properties are needed to causally account for paradigm physical effects, then closed physics will include mental properties, and the contrast between the physical a and the mental will be removed. Paradigmatic mental properties could causally seep through the macro-micro hierarchy and t-monism would not be false.

But Papineau considers this criticism to miss its mark, as it fails to take note of his *empirical* hypothesis that it is highly improbable that mental properties will *need* to be included in the set of properties required by closed physics. If the empirical hypothesis is false, t-monism indeed becomes trivial. But with the supplementation of the exclusion hypothesis, the t-monism that then follows is of a non-trivial kind. A non-trivial t-monist will understand physics in terms of closed physics, but in addition maintain that the empirical hypothesis is probably true. On the other hand, an interactive mentalist will maintain that in order for physics to provide a complete causal account of paradigm physical effects it *will* have to undergo a radical shift in the kind of properties that it appeals to, as the fact that physics does not invoke mental properties is precisely the reason why it is unable to causally account for all paradigm physical effects. Therefore the empirical hypothesis and thus non-trivial monism are *probably* false. Papineau thus shifts the debate from one of 'monism versus anti-monism', to one of 'non-trivial monism versus trivial monism'.

Papineau's definition of the 'physical' does resolve Hempel's dilemma. Unlike c-monism, t-monism provides a definition of the 'physical' from which its (probable) falsity does not follow. And although as with f-monism, t-monism refers to a physics not yet within our grasp, unlike f-monism one can be assured of t-monism's correctness. But does a slight variation of Hempel's dilemma arise at a deeper level of Papineau's account? Does the problem that the f-monist faced in trying to provide justification for the claim that complete

¹⁴ See Crane (1991a).

physics is (probably) causally closed, reappear when trying to justify non-trivial t-monism? To support f-monism one must be able to assess the probability of the hypothesis that complete physics will not need to appeal to non-physical a properties. To support non-trivial t-monism one must be able to assess the probability of the hypothesis that closed physics will not need to include mental properties. Can one do this?

Well, Papineau believes that one can assess the probability of the empirical hypothesis, considering it highly improbable that mental categories will need to be included within a closed physics. He accepts that the categories of current physics and closed physics will not be identical, as it is unlikely that a causal account of all paradigm physical effects can be achieved with only the categories of current physics. Nor does Papineau suggest that one can hope to know what these categories will be. However, Papineau is confident that to completely explain paradigm physical effects, physics will never need to be supplemented with mental categories.

It is important to observe the shift in the argument that Papineau's definition of the 'physical' allows. The f-monist has been concerned with the question of whether in order to causally account for effects within the physical a domain (where the physical a domain is identified with a complete physics) one need appeal to *any* non-physical a properties. Papineau has focused the discussion upon the particular sub-section of non-physical properties of relevance to the mentalism/ anti-mentalism debate. For the purpose of the mental causation debate, he is only attempting to show that whatever changes physics undergoes, the one fact that will not alter is that mental properties will not need to be included within closed physics. One could, of course, construct alternative empirical hypotheses which contrasted the physical a with other domains of properties, for example, the neurological or the chemical. Nor need one regard the various empirical hypotheses as equally plausible. One might regard it as highly probable that closed physics will not need to appeal to mental properties, but improbable that closed physics will not need to appeal to chemical properties. Hence, to deny interactive mentalism one does not need to know what properties a closed physics will include, nor does one need to know every property that it will not include. One only needs to provide plausible argument to suggest that it will not include mental properties.

In making this suggestion, Papineau has shifted the emphasis from that of showing that physics will never need to appeal to mental properties because it can causally account for all paradigm effects by appealing to properties of the kind that current physics would classify as physical a, to that of showing that physics will never need to appeal to mental

properties to account for all paradigm physical effects because mental properties are not the kind of property that physics would ever need to appeal to. Given Papineau's argument, the mental cause of a physical effect will still be causally overdetermining. But the risk of overdetermination is not based upon some consideration to suggest that a property of the kind recognised by current physics must be waiting in the wings. Rather, it is because it is highly plausible that there will be a non-mental property waiting in the wings, and this is entirely due to the direct consideration of the empirical improbability of closed physics ever needing to appeal to mental properties in order to explain paradigmatic physical effects.

However, to relativise one's argument in this way one need not trivialise the causal closure principle. The f-monist could maintain a causal closure principle according to which; 'Every paradigmatic physical effect has a direct sufficient non-mental cause'. If this is combined with the claim that mental causes have paradigm physical effects, and the denial of systematic causal-overdetermination, this means that mental causes must be identical with something that is non-mental.¹⁵ I would suggest that this is the preferable option, because, unlike the trivial understanding of the 'physical', it does not run contrary to the physicist's understanding of the 'physical'. Having seen how the monist might plausibly define the term 'physical' for the purpose of advancing a causal closure argument, let us now assess the plausibility of the various causal closure principles.

¹⁵ See Papineau (2000).

4

Empirical Support for 'No effect has a Non-Physical Cause'

4.1 An Energy Transference Theory of Causation

There is one theory of causation, the energy transference theory of causation, or more precisely one interpretation of this theory, that furthers the strongest causal closure claim, namely that 'No effect has a non-physical cause'. The energy transference theory of causation maintains that causation involves a transference of a quantity from cause to effect, where this quantity is to be identified with energy or momentum. It is maintained by among others, Aronson (1971), Fair (1979) and Hart (1988).

The energy transference theory is not motivated by considerations of conceptual restrictions upon the notion of causation, but is instead the result of an empirically discovered identity that causation is identical with energy and momentum transfer, hence, insofar as it offers support for the strong causal closure claim, it offers empirical support. However, it is important to distinguish between two different interpretations of the claim that causation is empirically identifiable with energy and momentum transfer. The first weaker understanding, is that the essence of causation is not energy or momentum transfer, but as a matter of contingent fact all causal relations happen to be accompanied by energy or momentum transfers. The identity of causation with energy and momentum transfers is a contingent one. Hence, for example, one may advance a counterfactual theory of causation, according to which C causes E if and only if E counterfactually depends upon C, but have empirically discovered that all such counterfactual dependencies happen to be accompanied by energy or momentum transfers. The alternative stronger claim, is that although the identification of causation with energy and momentum transfer is an a posteriori discovery, it is a necessary a posteriori truth, in the same kind of way that 'Water is H₂O' is a necessary a posteriori truth. Hence, it is in virtue of the fact that there is a transfer of energy or momentum between C and E that there is a

causal relation between C and E. This is a very important distinction, and it shall figure heavily in later discussion. If, however, one is advancing an energy transference theory of causation, one will interpret the claim that causation is identical with energy and momentum transfer in the stronger sense — causation, qua causation, consists in energy and momentum transfer.

Why maintain an energy transference theory of causation? Well, it certainly appears to offer a very persuasive account of causation, especially when compared with other theories of causation. After considering the serious problems facing many of the generalist accounts of the causation — the regularity theory's inability to distinguish between accidental coincidences and causal sequences, and its a priori closure of apparently empirical questions — the Neo-Humeans uninteresting distinction between causal and nomic regularity — the sheer number of ad hoc postulates that Lewis has had to tack on to his counterfactual analysis of causation in order to give an adequate account of causal direction, deal with epiphenomena and resolve the problem of redundant causes, one is left with the feeling that these generalist theories are missing out something of importance. In the light of this, singularist theories become all the more attractive. The move from generalist theories of causation to singularist theories of causation, is further supported by the recognition that the properties of causation are particulars, not universals. (As discussed in Part One, §4.6).

Of these singularist theories of causation, the energy transference theory appears to be both plausible and to remove many of the standard problems within causation. In many of the causal relations that our paradigm causal statements refer to, an energy transfer does take place. If the baseball causes the windows to shatter, then the kinetic energy of the scattered chunks of glass has its source in the kinetic energy of the ball, which is transferred on impact. If the sun causes the earth to warm, then energy of fusion is transferred by photons, becoming the thermal energy of terrestrial objects. Furthermore, as with all transference theories, the energy transference theory is able to explain how we come to use much of the causal language that we do. Transitive verbs such as 'push', 'pull', 'lift', 'knock', can all be replaced by the term 'cause'. As transitive verbs play such a central role in our causal language, a theory of causation should be able to accommodate this fact. All transference theories of causation can, as most (if not all) transitive verbs denote a process of transference. For example, assuming the energy transference theory of causation, 'Fred knocked the book on to the floor' is true, only if there is a transference of momentum from Fred to the book.

The energy transference theory is also consistent with the structural features of causation — distinctness, transitivity and asymmetry. The condition of distinctness follows as a consequence of a transference theory of causation. It is implicit within the notion of 'transference' that if a and c are identical, then a quantity cannot be *transferred* from a to c. Therefore in order for the transference of a quantity to take place between cause and effect, cause and effect must be distinct. Transference is also transitive; If A transfers p to B and B transfers p to C, then A transfers p to C. Finally, the asymmetry of energy transference provides an obvious explanation for the asymmetry of the causal relation. The direction of causation is established by the direction of the transfer of energy; energy is transferred from cause to effect, not from effect to cause. Consequently, whilst many theories of causation, particular those that are generalist, struggle to provide an account of causal priority in terms of temporal priority, and in doing so, typically rule out the possibility of simultaneous and backwards causation, a plausible account of causal priority follows from the core ideas of the transference theory. As it is able to give a sound account of causal priority, it does not encounter the problem of effects and of epiphenomena, faced by generalist theories such as Lewis' counterfactual theory and the naïve regularity theory.

The energy transference theory is also a theory of causation that is consistent with paradigm causal claims that most theories of causation find problematical. Most theories of causation have particular difficulty in reflecting our intuitions in cases of redundant causation. Such cases arise if there is more than one causal chain which would each have been sufficient to produce the effect, thus if causes are pre-empted or an effect is causally overdetermined. As an example of a pre-empted cause, imagine that Harriet and Edward are both shooting at tin cans, and Edward's shot hits the can first, causing it to fall over. Let us say that if Edward had not shot the can, it would have still fallen over because Harriet's shot would have hit it. While such cases of pre-emptive causation provide a huge stumbling block for counterfactual theories of causation, they are easily accommodated by the energy transference theory. Given this theory of causation, one can correctly affirm that it was Edward's shot that caused the can to fall over, not Harriet's. This is because there was a transfer of energy from Edward's bullet to the tin can, but not from Harriet's bullet to the tin can. Cases of causal overdetermination are also easier to accommodate within an energy transference theory. For example, take the example of two baseballs causing a window to shatter. Neither cause is individually necessary for the effect, as the window would have shattered even if just one baseball had hit it. The energy transference theory can allow, however, that the two baseballs cause the window to shatter, as each is a source of the energy that goes into the shattering. (Of course, more fine-grained cases

of causal overdetermination can be advanced which would pose problems for the energy transference theory, but in these cases one must ask whether there is a fact of the matter about what the cause is). Finally, many of the counterexamples against the energy transference theory — the most common of which point to supposed causal relations involving no energy transfer or an energy loss — can plausibly be explained away or rejected.¹⁶

4.2 Fair's Solution to the Problem of Hume's Similarity Circle

There is however one very serious problem with this theory of causation. The energy transference theory of causation tries to provide a theory of causation by showing that the word 'energy' has roughly the same meaning as 'cause'. An alternative theory approaching the problem from a similar direction could acknowledge the near synonymy between 'cause' and 'produce', and thus maintain that 'A causes B if and only if A produces B'. Hume advances a complaint against all proposals of this type:

'... the terms of efficacy, agency, power, force, energy, necessity, connexion, and productive quality, are all nearly synonymous; and therefore 'tis an absurdity to employ any of them in defining the rest' (1738, p.157).

In other words, causation cannot be understood in terms of energy transference, because the meaning of 'energy' is so heavily linked to the meaning of 'cause', that 'energy' cannot be defined in an informative way independently of 'cause'. To provide a non-circular definition of the term 'cause' one has to break out of this circle.

Aronson (1971), for example, provides no way of replying to Hume, because he does not expand on what he means by the term 'energy'. Others, such as Fair (1979), believe that they can resolve Hume's problem. Fair argues that if Hume is taking the words 'power', 'force' and 'energy' as understood in everyday language, then the criticism of synonymy and interdefinability stands. However, in physics the meaning of these terms has been sharpened. 'Power', 'force' and 'energy' all have explicit definitions in physics in terms of mass, position, wavelength and their time-derivative and thus all have senses in which

¹⁶ I refer the reader to Aronson (1971) and Fair (1979) for such responses.

they are not synonymous with the other terms in Hume's causation circle nor with each other. (Fair (1979, pp. 222-3)).

Although Fair's proposal allows one to break out of the Humean circle, it generates two further problems. The first is the less difficult to respond to. If causation is constituted by the transference of energy and momentum in the technical sense of physics, then how is it that people who know nothing about physics are able to accurately discriminate between causal and non-causal sequences? Given that 'Causation is energy transfer' is a necessary a posteriori truth, I would apply a Kripkean/ Putnamian account of the way in which natural kind terms enter our language to the term 'cause'. We have come to call those relations that exhibit, for example, counterfactual dependence, 'causal'. Subsequently, it has been empirically discovered that causation is identical with energy transfer. It is this that fixes the extension of the term 'cause.' Furthermore, this identification enables one to explain why one associates causation with, for example, counterfactual dependencies. Counterfactual dependencies often serve as evidence that there is a transfer of energy, because energy transference theories explain why causal relations tend to exhibit counterfactual dependence. For example, the striking causes the match to light. Without this striking, in possible worlds most similar to the actual world, the lighting does not occur. Why? Because in these worlds, as with the actual world, there will be no transfer of energy. But energy transfer and thus causation does not demand counterfactual dependence, as we find with the case of redundant causation. Thus one might say that an energy transference theory of causation provides the correct account of causation, and counterfactual dependence provides a reasonably reliable tool that can be used to identify energy transfers. Why does causation play such a central role in our language and life? Because it is important that we are able to pick out those sequences which involve energy transfer.

But given Fair's proposal one still faces a further problem that is very important to the mental causation debate and that cannot be satisfactorily addressed. This is that Fair's interpretation of the energy transference theory leaves no room for non-physical causation. In order to break out of the Humean circle, Fair identifies 'energy' with what he considers to be 'physical energy'. If Fair is correct, all causation involves transference of a *physical* quantity, and thus all causation reduces to physical causation. Hence the strong causal principle — 'No effect has a non-physical cause' follows from Fair's account. I would suggest that any theory of causation that implies that 'No effect has a non-physical cause' should be held in a highly dubious light.

Clearly, any theory of causation that entails the rejection of mental causation should be dismissed. This will be agreed by those on both sides of the mental causation debate, for it begs the question not only against the interactive mentalist, but the psychophysical reductionist, who maintains that there is mental causation, but identifies it with physical causation. Given Fair's interpretation of the energy transference theory of causation, one is not forced to deny that there is mental causation, but one is forced to deny that mental causation is non-physical causation. That is, to reconcile Fair's theory of causation with the claim that there is mental causation, one must first assume the correctness of psychophysical reductionism. But the interactive mentalist will argue that this is to beg the question against his position. Current physics does not allow one to decide upon the correctness of Fair's identification of causation with the transfer of physical energy. He must therefore appeal to the idea that a future completed physics, that is non-trivial, will enable one to account for all causation in terms of the transfer of physical energy. But whether or not such a physics could exist, is precisely the point at issue between the monist and the mentalist. Hence, Fair's theory of causation cannot be appealed to by the psychophysical reductionist in order to defend causal closure, for in assuming that all causation is physical causation, he is assuming the very thing that the mentalist is questioning.

One may respond that Fair's account of causation does not stem from mere monistic prejudice but a certain ranking of the desiderata for evaluating theories of causation. As we have seen the energy transference theory is able to address certain problems, dominant within the philosophical literature on causation, and this is more important than the fact that, in doing so, it links causation with underlying physical mechanisms. But I would question this ranking of the desiderata. Furthermore, an energy transference theory of causation certainly need not entail that all causation is physical causation. Fair's energy transference theory of causation entails that no effect has a non-physical cause because of two central premises within his account. The first premise is that energy is essentially physical, which as we have seen Fair appeals to, to enable him to break out of the Humean circle. The second premise is that all causation consists in energy transfer. I will go on to argue that from the consideration that the energy transference theory of causation provides a good account of purely physical causation, one cannot conclude that it can be appealed to provide an account of all causation.

Hence, in summary, if one assumes an energy transference theory of causation, that energy is essentially physical, and the homogeneity of causation, this entails that no effect has a non-physical cause. But to motivate this understanding of causation, is to beg the

question against the interactive mentalist, and for this reason, at least one of the premises must be rejected.

5

Empirical Support for 'No physical effect has a non-physical cause.'

5.1 The Conservation Laws of Physics

I shall now consider another kind of monistic argument for a strong causal closure principle, which also arguably requires the adoption of an energy transference theory of causation, or at the very least, some variant of it. A system is conservative if its total amount of energy and momentum can be redistributed, but not altered in amount, by changes that happen within it. The laws of conservation of energy and momentum state that every physical system is conservative or is part of a larger system that is conservative.

An appeal to the conservation laws of physics allows one to dismiss certain forms of interactive dualism. (Note that here physics is being identified with current physics and hence these laws are not beyond question, precisely because they are based within a physics that is most probably incomplete. However, for the purpose of this discussion I shall assume their correctness). In particular, Descartes' theory of psychophysical interactionism appears to conflict with the law of the conservation of momentum. Descartes' own law of energy conservation left a causal gap within the physical domain, which Leibniz (1898) suggested that Descartes thought that mental causes could fill. Descartes' mechanics held that 'quantities of motion' (by which he meant mass times speed) were conserved, but because Descartes' 'quantity of motion' was a non-directional quantity, the direction of a body's motion could be changed without a change in its quantity of motion. This meant that mental substances could causally affect physical substances by altering the direction of their motion without any violation of his conservation laws. Descartes therefore suggested that the mind altered the direction of the motion of particles in the pineal gland, these motions ultimately giving rise to bodily behaviour. However, Descartes failed to recognise that momentum was a vector. Because

momentum is a vector and because it is a conserved quantity, even if mental substances simply alter the direction of a moving physical substance this leads to a violation of the law of the conservation of momentum. So because of its proposal that the *non-physical affects* the physical by altering its direction of motion, Descartes' psychophysical interactionism conflicts with well-confirmed empirical conclusions within current physics, and hence given the assumption that these laws are correct, it should be rejected.¹⁷

But the conservation laws of physics are sometimes appealed to, to dismiss, not just Descartes' form of interactive mentalism, but interactive mentalism in general, the assumption being that if mental causes have physical effects, then unless mental causes are physical causes, this must conflict with either the law of the conservation of energy or the law of the conservation of momentum. For example, as Fodor enquires; 'how can the non-physical give rise to the physical without violating the laws of conservation of mass, of energy and of momentum?' (Fodor, 1994) Indeed, I consider Papineau's (2000) comments about the conservation laws of physics to echo a general sentiment amongst many of those who support a closure principle:

'When I first became interested in the causal argument a few years ago, I recognized that there were many points where it could be queried. However, I assumed that the completeness premise itself was quite uncontroversial.... My original thought was the completeness of physical would follow from the fact that physics can be formulated in terms of conservation laws. If the laws of mechanics tell us that important physical quantities are conserved whatever happens, then doesn't it follow that the later states of physical systems are always fully determined by the earlier states?' (2000, pp. 184-5).

It is, note, the very possibility of psychophysical causation which appears to conflict with the conservation laws of physics, the question being how can the mental, if it is non-physical, make a causal difference within the physical domain without violating these laws? Hence an appeal to the conservation laws of physics allows one to advance the strong causal closure claim that; 'No physical effect has a non-physical cause'.¹⁸ To appeal to the conservation laws of physics to dismiss interactive mentalism is therefore to advance a strong causal closure argument in which the premise of the denial of systematic causal overdetermination is rendered redundant.

¹⁷ For a discussion of these points see Woolhouse (1985).

¹⁸ For a defence of this point see Lowe (1996, Chap 3) and Lowe (2000).

5.2 The Premise that Energy is Essentially Physical

But in fact, on their own, the conservation laws of physics do not entail that no physical effect has a non-physical cause. In order for them to do so, two further premises must first be assumed. The first is that energy is essentially physical.

If there is a 'psychic energy' that operates conservatively, this is perfectly consistent with the conservation laws of physics, for the conservation laws of physics certainly do not tell us which fundamental forces there are, only that they must operate conservatively. It is consistent with current physics that in addition to the physical forces of gravitation, electromagnetism and strong and weak nuclear forces, psychic force may be a further fundamental force. This is Hart's (1988) substance dualist response to the problem of mental causation. He proposes that there is such a thing as 'psychic' energy which can be transferred along causal chains and which behaves in accordance with the conservation laws. Mental states cause physical states by transferring psychic energy, and psychic energy is converted into physical energy in accordance with the conservation laws. (Given a property emergentism, psychic force may be identified with the 'configurational forces' that McLaughlin (1992) discusses. These are forces which only arise at higher levels on the macro-micro hierarchy).¹⁹

Although the conservation laws of physics are wholly consistent with the possibility that there is psychic energy, two kinds of argument might be presented to suggest that it is implausible that psychic energy actually does exist, or, at least, that if it does, that it could be non-physical. Let us start with the argument against the existence of psychic energy as a form of non-physical energy. The monist might argue that if psychic energy was a product of a further type of fundamental force, then it would have to be a new form of physical energy. The concept of energy is so intimately tied to that which is physical, that one could not have such a thing as non-physical energy. Energy is essentially physical.²⁰ Hence, if there is a further type of energy that can be appealed to in order to explain psychophysical interactions, it is just a further type of physical energy. And, of course, to identify psychic energy with physical energy is to reduce psychophysical causation to physical causation and thus advance a monism. Hart hints at this sort of possible criticism

¹⁹ Note that McLaughlin, himself, does not consider such configurational forces to exist.

²⁰ See Lowe (1996, p. 61) who considers and rejects this argument.

when he suggests that 'although we would want psychic energy to be specified from the beginning in psychological terms, the orthodoxy might feel ill at ease calling such energies as pass through the central nervous system psychic, and the orthodox would probably expect such energy to be electrochemical' (1988, p. 133). But his response to this sort of charge should be to ask why 'energy' must be physical energy, for it is not a consequence of current physics. Of course, the monist may define all energy as 'physical energy'. Horgan, for example, comes close to doing this. In his statement of the closure principle he assumes that 'all fundamental causal forces are physical forces' (1993, p.560). But this is to trivialise one's definition of the term 'physical' and thus to trivialise monism. Hart must consider, in a non-question begging way, what features unite physical energy and whether or not psychic energy possesses these features. Certainly, if psychic energy possesses features that are radically different from all of the various types of physical energy one would have good reason to resist referring to it as physical. Of course, one could call this new kind of energy 'physical' but then one would have extended the notion of what it is to be 'physical'.²¹

The second argument against the existence of psychic energy is to simply point to the fact that that we have no direct evidence for the existence of this non-physical form of energy, hence no reason to think that it actually exists. Despite all our knowledge about biochemical and neurophysiological processes, none gives us any evidence to suggest that psychic energy exists. Hence, Papineau argues that; 'If there were such special forces, they could be expected to display some manifestation of their presence. But detailed physiological investigation failed to uncover evidence of anything except familiar physical forces' (2000, p. 202). Hart acknowledges this fact: 'We do not claim to have shown that there exists such a quantity as psychic energy. To support such a claim, it would be necessary to specify (in, one would expect, purely psychological, though perhaps now unknown, terms) a quantity conserved through and traceable along (almost) all naively identified wholly intrapsychic causal chains. We have specified no such quantity and are, alas, unable to do so; we do not know whether there is such a quantity' (1988, p. 130). But in response to this criticism one can make two points. Firstly, despite the fact that there is no evidence to suggest that psychic energy exists, these considerations reveal that one can consistently maintain an interactive mentalism, without denying the conservation laws of physics. Secondly, someone such as Hart would not be persuaded

²¹ I suppose if one was worried about using the term 'energy', one could instead appeal to Castaneda's (1984) 'causity'. According to Castaneda 'causity' is a quantity that is transferred from cause to effect. Causity, is, as a matter of fact energy, but this is a contingent identity. Unfortunately, Castaneda does not say enough about the notion of 'causity' to show how it could be anything other than energy.

that physical energy could account for psychophysical interactions in a satisfactory way, and hence until the existence of psychic energy is recognised, a gap will remain in our causal account of the physical domain. To this the monist will respond that it is implausible that psychic energy will be needed to fill any such gaps — a point to which I shall return later.

5.3 The Premise of the Energy Transference Theory of Causation

Even if, as a matter of fact, all energy is physical energy, the conservation laws of physics still do not entail that no physical effect has a non-physical cause. This would only follow if the only way that a physical system could be causally affected was by affecting the quantity of energy within it or by affecting its momentum. This does appear to be the assumption of some within the mental causation debate. Hence consider a quote from Searle:

'In order for us to have radical freedom, it looks as if we would have to postulate that inside each of us was a self that was capable of interfering with the causal order of nature. That is, it looks as if we would have to contain some entity that was capable of making molecules swerve from their paths. I don't know if such a view is even intelligible, but it's certainly not consistent with what we know about how the world works from physics' (Searle (1984 p. 92)).

Searle's suggestion is that in order for mental causes to have physical effects, mental causes would have to make molecules change their direction thus violating conservation laws. Searle is thus assuming that the only way that the mental could causally affect the physical is by altering the motion of physical substances, rather in the way that Descartes proposed. Can one make this assumption?

More accurately, given that the conservation laws hold not only in cases of purely physical interactions but also in any interaction in which the physical is involved, then if all energy is physical the conservation laws imply that (1) Non-physical causes cannot affect the quantity of energy within a physical system and (2) Non-physical causes cannot affect the

momentum of a physical system. The question is does the combination of (1) and (2) entail that 'No physical effect has a non-physical cause'?

I would suggest that how easy one finds this question to answer, will depend upon the theory of causation that one adopts. In particular, if one assumes an energy transference theory of causation, then (1) and (2) will entail that no physical effect has a non-physical cause, because according to this theory of causation all that there is to a cause having an effect, is the transferring of energy from cause to effect.

It therefore comes as no surprise that Hart's theory of interactive dualism assumes an energy transference theory of causation. (See Hart (1988, ch. 5)). According to him all causation consists in energy transfer. For this reason, in order to maintain an interactive dualism, he is forced to claim that physical energy is only one type of energy, hence allowing him to advance the claim that mental causation, although involving energy transfer, does not involve the transference of physical energy. Hart's response to the problem of mental causation is the only one available for the interactive mentalist, given the theory of causation in which Hart embeds psychophysical causation.

And indeed, returning to Descartes' interactionism, it is interesting to note that Descartes also arguably maintained a kind of transference theory of causation in which the quantity that was transferred was not energy, but motion. According to this theory of causation, change in the physical world reduces to transfers of motion, for any change in a physical body is, according to Descartes, due to matter (corpuscles) in motion that leads them in to contact with one another.²² But Descartes did not think that this motion transference theory of causation should be extended to causation involving mental states. For, as we have seen, Descartes did not allow the mental to cause the physical by affecting quantities of motion — this would have been incompatible with Descartes' conservation laws of physics, as according to them, an effect can gain only as much quantity of motion as its cause gives up. He instead suggested that the mental could alter the direction of a body's motion. Hence, Descartes is advancing a transference theory of motion in the case of purely physical causation, but considers this model of causation to be inappropriate to psychophysical and psychic causation. He is, that is to say, denying the homogeneity of causation. From the perspective of current physics, we can see that Descartes' denial of the homogeneity of causation does not allow him to secure psychophysical causation, for

²² For an excellent discussion of the claim that Descartes' should be interpreted as maintaining a transference theory of causation, see Clatterbaugh (1999, ch. 2, pp. 28 ff.).

the causal role that Descartes allows the mental within the physical domain violates the conservation laws of physics.

To summarise the arguments that have been presented so far; the claim that no effect has a non-physical cause is entailed if one assumes 1) Energy is essentially physical 2) An energy transference theory of causation and 3) The homogeneity of causation. The claim that no physical effect has a non-physical cause follows from the conservation laws of physics if one assumes 1) Energy is essentially physical and 2) An energy transference theory of causation that holds in all cases of physical and psychophysical causation. Note therefore, that this proposal is consistent with the energy transference theory having a narrower application than the one that Fair suggests. According to Fair all causation consists in energy transfer, whilst this proposal is consistent with the possibility that purely psychological causation does not consist in energy transfer.

If one rejects an energy transference theory of causation (at least in the case of psychophysical causal interactions), is the interactive mentalist still faced with the same limited option of maintaining either that there is such a thing as psychic energy that operates conservatively or adopting a theory of psychophysical causation that violates the conservation laws of physics? That is, does it remain true that a physical system can only be causally affected by affecting the quantity of energy within it or its momentum? Well, if one considers the weaker interpretation of the claim that causation is empirically identifiable with energy transfer, one can see that the monist may consistently reject an energy transference theory of causation whilst at the same time maintaining that as a matter of contingent fact all causal relations involving physical effects happen to involve energy transfer. Causation, qua causation, does not consist in energy transfer, but all relations involving physical effects, as a matter of fact involve energy transfers. However, as shall become clear, once one has moved away from an energy transference theory of causation, the claim that all causation involving physical effects involves energy transfer is much harder to sustain.

6

Empirical Support for Weak Causal Closure Principles

6.1 The Problem with Hart's Theory of Psychophysical Interactionism

Hart maintains an energy transference theory of causation, along with the homogeneity of causation. However, he is able to reject both of the strong causal closure principles because he rejects the premise that energy is essentially physical.

One might question whether Hart is able to advance a plausible energy transference theory of causation. Fair was able to break out of the Humean circle because he identified energy with physical energy. This is clearly not an option for Hart. Hence the problem is whether or not he can provide an informative understanding of the term 'energy'.

There is, however, a further reason why the interactive mentalist should resist Hart's solution to the mental causation debate. Let us say that M1 is your decision to raise your arm; B1 is your arm rising; N1 and N2 are intervening states in your nervous system. Hence we have the following causal chain of states:

$$M1 \rightarrow N1 \rightarrow N2 \rightarrow B1$$

According to Hart, the causal relation between each of these states consists in a transfer of energy; each cause increases the quantity of energy or momentum of its effect. M1 is needed within the causal chain of states to explain the increase in the quantity of energy or momentum of N1. The problem is that the monist will respond to Hart's proposal by asking why we should think that there is any such gap in the chain of transfers of energy. Even if current physics cannot equip us with an account of what it is that causes the

quantity of energy or momentum of N1 to increase, it is probable that future, completed physics will. There is, in other words, no need to appeal to psychic energy because it is improbable that a completed physics will leave any gaps in the chains of energy transfer.

Thus, yes, an appeal to psychic energy enables us to reject what I consider to be the 'strongest' arguments for strong causal closure principles. However, the weaker kind of causal closure principle, appealed to within the argument from causal overdetermination, remains. Hence, as discussed earlier, Papineau (1990, 1993) advances a definition of the physical into which this kind of causal closure principle is built. He supplements this trivially true causal closure principle with the empirical hypothesis that 'closed physics' will never need to appeal to mental properties in order to explain paradigm physical effects. If the empirical hypothesis turns out to be true, then a non-trivial monism is correct. In response to Hart's theory of interactive mentalism, Papineau would probably respond that it is highly implausible that there will be any gaps in the chains of energy transfer that can only be filled by psychic energy. According to Papineau, current physics 'aims to develop a complete theory of paradigm physical effects in terms of the categories of energy, field, and space-time structure' (1993, p. 31). Admittedly, it is highly unlikely that current physics is complete and wholly accurate, hence it is highly unlikely that current physics leaves no gaps in the chains of energy transfer, or that it always fills them in the correct way. But although the categories of current physics will probably need to be supplemented before these gaps can be filled, it is implausible that physics will ever need to appeal to psychic energy in order to do so.

To this the interactive mentalist could respond by rehearsing the various arguments in favour of the existence of psychic energy — there is nothing in physics to rule out the existence of psychic energy, and until one does appeal to psychic energy causal gaps will remain. But an interactive mentalist should not get into this debate. He should in fact reject the energy transference theory of causation when applied to psychophysical causal interactions.

An interactive mentalist who assumes an energy transference theory of causation invites the argument that mental causes will never need to be appealed to within the physical domain, because the energy transference theory of causation is biased towards a physical model of causation. It is a theory of causation that is generated by examining cases of purely physical causal interactions and then applying it to all cases of causation. The examples that Fair (1979), Aronson (1971) and even Hart (1988) give to motivate the energy transference theory of causation are all purely physical examples. The fact that the

baseball causing the window to shatter, John's body causing the door to move, the sun causing the earth to warm can all be explained in terms of energy transfer, shows, at best, that many of the causal relations to which our paradigm physical causal statements refer, can be understood in terms of energy transference.²³ It does not show that all causation can be understood in terms of energy transference. This leads me to make two points.

Hart is attempting to make sense of psychophysical causation on the basis of a model of causation designed to explain purely physical causation. He provides support for the energy transference theory of causation by considering cases of purely physical causation, and then attempts to fit psychophysical causation and psychological causation in with this account. He shows that an energy transference theory of causation may be made compatible with interactive mentalism by appealing to psychic energy. But in doing so, he makes the mental behave causally in a similar way to the physical. And precisely because he proposes a model of psychophysical causation in which mental causation is made to imitate physical causation, it is no wonder that it seems likely that any possible gap in the chains of energy transfer that mental causes could fill within the physical domain, could be more appropriately filled by physical causes.

Secondly, at best, those who advance the energy transference theory of causation show that this theory of causation provides a good account of physical causation. They do not show that the energy transference theory of causation is the most appropriate model of causation for understanding non-physical causation. I fail to see why one should try to make psychophysical causation fit in with accounts such as the energy transference theory of causation, because I fail to see why one should try to make sense of all causation, in the light of what is known about purely physical causation. Hence, even if the causal claim that 'Every physical effect has a set of direct physical causes which together are sufficient for its occurrence' seems plausible if one assumes an energy transference theory of causation, this should not lead one to adopt monism.

These considerations become particularly interesting when it is recognised that, contrary to their indications, monists often make assumptions about causation in general, and psychophysical causation in particular, that are most compatible with an energy transference theory of causation. These assumptions make the problem of mental causation look like a very hard one for the interactive mentalist to solve.

²³ See Fair (1979, p. 229) for these examples.

6.2 The Energy Transference Theory of Causation & Assumptions within the Mental Causation Debate

Consider Hart's substance dualism. The form of psychophysical interactionism that Hart proposes leads to various difficult questions. How could there be an energy flow between a disembodied mind and matter? Indeed, given Hart's position it is just as hard to explain how the mental could be affected by the physical as it is to explain how the physical could be affected by the mental. Is it conceivable that electro-magnetic energy could be converted into a quantity of belief and vice versa, or are they too dissimilar for such a conversion to take place, thus leading one to question psychophysical causation? Hart considers the fact that the question of psychophysical causation is particularly problematical given the energy transference theory, to be one of the main virtues of this theory of causation. The energy transference theory of causation makes the problem of mental causation 'properly difficult' (1988, p. 66). I, on the other hand, consider the energy transference theory to create unnecessary barriers in the mental causation debate.

To see exactly why the energy transference theory of causation makes psychophysical interactionism so difficult, let us consider some of the arguments commonly raised against interactive mentalism. It is rarely stated which theory of causation is being appealed to within the mental causation debate, perhaps because it is considered that the claims that are being made are neutral amongst the various theories of causation. Those that do consider what it means to say that 'a is a cause of b', gesture towards a nomological theory or perhaps a counterfactual theory. But in fact many of the implicit assumptions made within the mental causation debate (invariably by the monists) about causation in general, and psychophysical causation in particular, are only compatible with a far more rigid theory of causation; one that is most accurately represented by an energy transference theory.

In the first place, philosophers of mind often assume that a causal mechanism must be provided in cases of psychophysical causation. For example, Smith and Jones (1986) are under the impression that causation requires mechanisms:

'If we say of two events that the first caused the second, then the question "*how* did the one cause the other?" is usually in order...in claiming that

there is a real causal relation here, we do commit ourselves to there being some linking mechanism or other...' (p. 53)

From the above, in asking for a causal mechanism, what is being asked for is an explanation of how a cause is efficacious. If one cannot provide an acceptable account of how a particular cause brings about its effect, then this causal relation should be rejected. Importantly, this is true not only in cases of indirect causation (where one could simply appeal to the intermediary state in order to explain how a cause brings about its effect) but also in cases of direct causation. That is, if one is committed to the idea of causal mechanisms proper, then one will require an answer to the question of how a cause brings about its direct effect.

This requirement makes the problem of mental causation look like a very difficult one. Even if the interactive dualist is able to show that he can offer a response to the argument from causal overdetermination, he is still required to explain how a mental cause could have a physical effect. Hence, for example, if one maintains a substance dualism in which mental substances are non-spatial, one of the standard objections is that one cannot explain how a non-spatial substance could causally affect a spatially located substance.

But demanding a causal mechanism also makes purely physical causation difficult to understand. Even within fundamental physics one cannot find connections that would enable one to explain how one physical state caused another. Smith and Jones' (1986) recognise that beyond a certain point, there is no asking how. Eventually one will reach the rock bottom level at which there is no further explanation to be given. Unfortunately, they consider that cases of psychophysical causation are precisely the sort that cry out for a causal mechanism — 'an immaterial cause and a physical upshot are even less like each other.... so in this case the question 'how does the causal mechanism work?' seems even more urgent' (p.53). Their conclusion — 'given that causation requires the existence of causal mechanisms, and that there can be no such mechanisms linking across the body/Mind divide, it follows that there can after all be no causal interaction between physical bodies and immaterial Minds' (p.54). Hence, Smith and Jones' suggestion seems to be that some cases of direct causation require a causal mechanism, and psychophysical causation is one such case.

Smith and Jones simply assume that all plausible theories of causation should enable one to understand the notion of a cause's efficacy, or, in other words, the fact that a cause in some way 'brings about', or 'produces' its effect. But this is not the case. According to

most theories of causation the search for a causal mechanism is in fact misguided. Although all would accept that there is nothing wrong in asking how a cause brings about its indirect effect, they cannot make sense of the question of how a cause brings about its direct effect. Thus generalist theories of causation, such as counterfactual theories, and naïve regularity theories all stop short of identifying any mechanism of causation. A regularity theory, for example, maintains that c causes e if c and e are constantly conjoined. There is no further fact about the cause or effect, or about the sequence $c \rightarrow e$ that gives it its causal character. Thus the question of what it is about a constantly conjoined sequence which explains how a cause is able to bring about its effect simply has no application. The same is equally true of non-reductive singularist theories of causation.

The one group of theories of causation that do consider the question of how a cause brings about its direct effect to be legitimate, are the reductive singularist theories — the category into which the energy transference theory of causation falls. The reductive singularist analyses causation in terms of the underlying non-causal processes associated with causation. Consequently, he can appeal to these non-causal processes in order to explain how a cause brings about its direct effect. For example, if one assumes an energy transference theory of causation, in response to the question of how a cause brings about its direct effect, one may reply that it is by transferring energy to its effect. For this reason, it is perfectly legitimate to require an explanation of how a mental cause brings about a physical effect, given Hart's theory of mental causation.

Now of those theories that allow the question of how a cause brings about its direct effect — the reductive singularists' — not all are biased towards a physical account of this mechanism. That is, the singularist noncausal process that connects a cause to its effect, and to which one appeals in order to explain how a cause brings about its direct effect may not be a process that is biased towards a physical interpretation. For example, Ehring's (1997) theory of causation as trope persistence is a reductive singularist position that identifies the causal process with trope persistence or partial persistence. Ehring's theory of causation allows the question of how a cause brings about its direct effect (it is by trope persistence), but this is not biased towards a physical understanding of this mechanism. Hence, the requirement of mechanism proper is distinct from the specific form of mechanism assumed within an energy transference theory of causation.

However, physicalists tend to demand a causal mechanism that is most appropriate to understanding the relevant process as a physical one, hence making their commitment to

an energy transference theory apparent. As Burge notes, in discussions of mental causation it is sometimes suggested that for mental causes to have effects, the mental must 'provide an extra bump on the effect' (1993, p.115). And furthermore, as Crane argues, many physicalists 'tend to talk in terms of causes requiring 'forces' or 'oomph' (1995, p. 231). He cites Horgan (1993) as one such example. Horgan, in discussing mentalism, explains that he uses 'the term 'efficacy' rather than 'relevance' because the latter seems too weak to capture the kind of *oomph* that higher-order properties ought to have if they are not epiphenomenal' (1993, p. 572 fn. 19). The notion that a cause must exhibit some kind of 'oomph' suggests that some sort of force must be in play. This is illustrated by a later comment of Horgan's, when in defining the completeness of physics, he explains that 'This means that non-physical properties cannot be causally basic properties — ones that generate fundamental forces that combine with physical forces to yield net forces different from the net resultants of physical forces' (p. 573).

Few causal theories require that causally efficacious properties must exhibit some kind of 'oomph', or involve bumping and impacting. According to most, bumping and impacting merely provide an instance in which one can confidently assert that there is a causal relation in existence. Causal relations, are not, however to be analysed in terms of bumping and impacting. It may be the case that all causal relations involve bumping and impacting, but this is to make a very different point. It is not to suggest that causal relations are causal because they involve bumping and impacting. An energy transference theory of causation, on the other hand, advances a model of causation in which the notion of a force plays a central role.

And, of course, as Crane (1995, p. 231) suggests, an understanding of causation in terms of forces, or bumping and impacting, makes the question of how there can be non-physical causation particularly difficult, for forces, and bumping and impacting are typically associated with that which is physical. This point is taken to its extreme, when Searle sarcastically comments: 'How could something mental make a difference? Are we supposed to think that our thoughts and feelings can somehow produce chemical effects on our brains and the rest of our nervous system? How could such a thing occur? Are we supposed to think that thoughts can wrap themselves around the axons or shake the dendrites or sneak inside the cell wall and attack the cell nucleus' (1984, p.17).

Hence, the interactive mentalist may propose that there could be non-physical forces, but in doing so it should be clear that he is adopting a model of causation most suited to analysing physical relations. Consequently, he encounters problems responding to weak

causal closure principles. For of course Papineau's empirical hypothesis appears plausible, if the 'gaps' in the determination of certain physical effects must be filled by the mental in the same way that the physical fills gaps, if mental states must behave like forces, if one must be able to explain the mechanism behind such relations, and furthermore these explanations must be physically respectable. But if one was to maintain a theory of causation that was not biased towards such a physical model of causation, would the same be equally true?

6.3 The Rejection of an Energy Transference Theory of Causation when Applied to Psychophysical Interactions

To assume an energy transference theory of causation in the case of psychophysical causation is to beg the question against the interactive mentalist because it is a theory of causation that is biased towards a physical model of causation. Psychophysical causal interactions, qua causal interaction, do not involve energy transfers. Consequently, either the energy transference theory of causation must be denied, or, if one is persuaded that the energy transference theory of causation provides a plausible account of purely physical causation, then one must deny the homogeneity of causation. (Note that this proposal is neutral within the mental causation debate. The set of entities that exhibit this alternative kind of causal relation may be an empty one). At best the energy transference theory of causation is a theory about only one kind of causation — physical causation. It is applicable only to the domain of purely physical causal interactions.

This is of direct relevance not only to the strong causal closure arguments, but also the argument from causal overdetermination because what it is to be a 'cause' and what it is to be an 'effect', and thus what it is to be a 'sufficient direct cause of an effect' depends upon one's theory of causation. If one advances an energy transference theory of causation and if one can give a sufficient physical account of every transference of energy that occurs within the physical domain, then it would be correct to conclude that every physical effect has a sufficient physical cause. But if there is more to the causing of a particular physical state than the transferring of energy to it, then one will not necessarily have identified a sufficient cause of the state simply by citing the thing from which the energy has been transferred.

Of course, one may argue that one can give a complete account of all change within the physical world in terms of energy transfer, whilst at the same time denying that one's account is causal in virtue of the fact that it involves energy transfer. Thus for example, one could maintain a counterfactual theory of causation while at the same time claiming that as a matter of fact physics can adequately account for every counterfactual dependency in terms of energy transfer. However, the monist must show that all causal effects upon the physical domain do in fact involve energy transfer, and the more one draws away from the energy transference theory of causation the harder this will be to do. From the perspective of an energy transference theory of causation, contrary to Hart, it seems improbable that gaps within our causal account of paradigm physical effects will ever need to be filled by mental causes, because it seems improbable that there are gaps in the chains of energy transfer that could not be filled by an appeal to physical energy. However, what it is for there to be a 'causal gap' in one's account, depends upon the theory of causation that one is assuming. It is possible that other, less restrictive, theories of causation will recognise causal gaps in the physical domain even if there are no gaps in the chains of energy transfer.

To explain what I mean by this let me give two examples. I hasten to point out that these examples only serve as analogies to help to show that what one sees as a 'gap' in one's causal account, depends upon one's causal perspective. Firstly, let us take Salmon's (1984, 1994) example of a pseudo-process. The rotation of a spotlight in the middle of a circular room results in a moving spot of light along the walls. According to Salmon, the moving spot of light is not a causal process because Salmon (who maintains a transference theory of causation, although not an energy transference theory) maintains that causal processes have a capacity to carry 'marks', but if a 'mark' is made in the moving spot at one moment it is not transmitted to later stages of the process. Any plausible theory of causation must allow one to draw the same conclusion as Salmon does — the moving spot is not a causal process. However, if one simply observes the moving spot of light one observes no obvious gaps, or at least no gaps of a spatio-temporal kind. Despite this fact, one would be wrong to therefore conclude that there are no causal gaps in such a process.

The energy transference theory of causation would obviously dismiss any such process as causal because energy is not transferred from the spot of light at time t_1 to the spot of light at time t_2 . However, according to the energy transference theory of causation, the only causes there are, are those that transfer energy or momentum, and the only effects there are are those that gain energy or momentum. For there to be a causal gap, given

the energy transference theory of causation, there must be a gap in the transfer of energy. Hence, it is only if one state had an energy gain and no physical state had a corresponding energy loss, that according to the energy transference theory, the physical domain would be causally open.

But other theories of causation may recognise 'causal gaps', whilst from the perspective of the energy transference theory there appears to be a seamless chain of causation. To give an example, imagine two people playing chess. If we simply consider the movement of the chess pieces, then a complete causal account of these movements could be given in terms of energy transfers. But is it the case that to give a complete account of the movements of the chess pieces, is to give a complete causal account of what is going on? What is not explained is why certain pieces are moved, why they are moved in certain directions etc.²⁴ It does not include the consideration that the movements of the chess pieces are directed towards a goal, that there are reasons for moving the chess pieces in certain directions rather than others. Neither does it seem obviously true that one could capture such considerations simply by appealing to energy transfer.

Alternatively, to give an example that Lowe (1996) advances to forward his theory of interactive dualism, imagine the movements of a spider on a web. If one was to ignore the web, and just look at the movements of the spider, we would observe no gaps in the spider's movement. However, it would be wrong to conclude that the web had no causal role to play in the spider's movements. The web 'enables and constrains these movements to take place in certain directions rather than others' (1996, p. 637). I think that it is fair to say that, at the very least, it is certainly not obvious that this additional causal role can be accounted for by the energy transference theory.

To summarise, it may be true that the physical world is causally-x closed, but it is only if all causation within the physical domain is of kind x that one should be led to identify mental causes with physical causes. If the monist understands x-causation in terms of transference of energy, and energy is identified with physical energy as Fair maintains, then the homogeneity of causation must be denied. Even if x-causation is understood in terms of transference of energy, and energy is not interpreted in such a narrow way, then unless one is to implausibly assume a model of mental causation based upon a model of physical causation, the homogeneity of causation must still be denied. In denying the homogeneity of causation, the monist must respond by advancing a strengthened closure

²⁴ This kind of point is also made by Broad (1929, pp. 108-109).

claim which allows that the physical world is not merely causally x closed. Alternatively, if the monist rejects the heterogeneity of causation, to avoid begging the question against the mentalist, he must not maintain an energy transference theory of causation within the physical domain. Causal relations need not involve energy transfers. If a causal relation does involve an energy transfer then it is not causal in virtue of the fact that it involves an energy transfer. Whether or not it is an empirical fact that all causal relations involving physical states involve an energy transfer is a further question.

6.4 Working Towards A More Plausible Model of Causation

What we need to do now is consider whether the claim that 'Every physical effect has a set of direct physical causes which together are sufficient for its occurrence' is plausible if one is not applying an energy transference theory of causation to psychophysical causation interactions.

If the energy transference theory of causation is rejected, then to what theory of causation should one turn? Alternatively, if, one does not wish to totally abandon the energy transference theory of causation, considering it to offer the most plausible account of purely physical causal interactions, and hence one denies the homogeneity of causation, one must ask how one should understand psychophysical and psychological causal interactions. Note that if the homogeneity of causation cannot plausibly be denied, the wrong conclusion to draw would be that there is no non-physical causation or that non-physical causation must be fitted onto this physical model of causation. If a satisfactory way of denying the homogeneity of causation cannot be developed in combination with the energy transference of causation, one must reject the energy transference theory of causation.

1. Regularity theories and counterfactual theories

Whilst energy transference theories make the problem of mental causation look very difficult, other weak theories of causation work in the opposite way. Hence, consider the naïve regularity theory and Lewis' counterfactual theory of causation — two theories of causation that I have suggested are implausible given that the properties of causation are CPs. It is often argued that the fact that there is thought to be a problem of mental

causation shows that these theories of causation are not being adopted. For example, to quote Heil and Robb: 'We take for granted that causal relations are something more than counterfactual dependencies or Humean regularities. Was causation nothing more than regularity or counterfactual dependence, the problem of mental causation — and more generally the mind-body problem — would subside.' (forthcoming, p. 16)

Does a regularity theory or counterfactual theory really allow the problem of mental causation to subside? Well, it leads to the rejection of strong causal closure principles, for then there is certainly nothing mysterious or problematical about mental states causing physical states. First, consider a naïve regularity theory. There are many systematic regularities between mental and physical states. Given the absence of any impediment, my desire to drink will be coupled with the act of drinking, my touching something very hot will be coupled with the feeling of pain and hence the withdrawal of my hand. If all there is to a mental state causing a physical state, is a mental state being constantly conjoined with a physical state, then there is no reason to think that there could not be psychophysical causation. Equally, the causal efficacy of the mental within the physical domain follows naturally from a counterfactual understanding of causation. 'My desire for water (qua being non-physical) causes me to pour a glass of water' is true, because possible worlds in which I have no desire for water and I do not pour myself a glass of water, are more similar to the actual world, than possible worlds in which I do not have a desire for water but I do pour myself a glass of water. Certainly, that my behaviour would have been the same in the absence of this mental state is incorrect on the most natural of interpretations.

Hence, the strong causal closure principles that 'No effect has a non-physical cause' and 'No physical effect has a non-physical cause' are certainly not supported by a weak theory of causation such as the naïve regularity theory or a counterfactual theory. Paradigm examples of non-physical property-instantiations are constantly conjoined with physical property-instantiations, and in many cases the latter do counterfactually depend upon the former. But even given these theories of causation, it is arguable that a problem of mental causation remains. The role of the causal closure principle within the argument from causal overdetermination is not to rule out the possibility of psychophysical causation. Its claim is simply that physics will never need to appeal to mental causes and it is the threat of systematic causal overdetermination that ultimately leads us to identify mental causes with physical causes. Thus for example, given a regularity theory, the monist will argue that although there is a relation of constant conjunction between a mental state M1 and a

physical state P1, the latter is also constantly conjoined with a physical state P2 and hence, unless we identify M1 with P2, P1 is causally overdetermined.

To this one could respond that given these theories of causation — the counterfactual theory, the regularity theory — what to identify as 'the cause' in cases of causal overdetermination is a notoriously difficult question, and that if a mental state and a physical state causally overdetermine an effect given such theories of causation, it is not unreasonable to argue that both are properly causes of the effect. But this would be to miss the point. It is the threat of systematic causal overdetermination that is the concern within the argument from causal overdetermination. That is, it is the appeal to the implausibility of the claim that as a general rule, mental causes have physical effects by causally overdetermining physical causes, that leads us to identify mental causes with physical causes.

I would suggest, however, that if one appeals to the argument from causal overdetermination and assumes either a regularity theory or a counterfactual theory, the plausibility of the causal closure principle is significantly weakened. That is, the claim that physics does not need mental causes in order to be causally complete is harder to support if all there is to causation is constant conjunction or counterfactual dependency. The monist will be required to show that, for example, for every regularity between a mental state and a physical state, there is also a regularity between this physical state and another physical state. It stands to reason that this will be harder to support empirically than those theories of causation that entail a narrower understanding of what it is for one thing to be causally related to another, and hence recognise fewer causes within the physical domain.

However, any position within the mental causation debate that assumes a counterfactual theory or regularity theory of causation in order to address the problem of mental causation, should be rejected. This is because both of these theories of causation offer implausible understandings of the causal relation, notoriously misidentifying the non-causal for the causal. Furthermore, they do not fit well with the claim that the properties of causation are CPs. Neither will it do to maintain that all physical causation consists in energy transfer, and that all non-physical causation consists in either counterfactual dependence or constant conjunction, for counterfactual theories of causation and regularity theories are just as problematical when they are applied to the non-physical

domain as they are when applied to the physical domain.²⁵ Indeed, the fact that these theories leave something out is confirmed when we consider psychophysical causation. We do not understand ourselves to be agents for broadly Humean reasons, or because mental and physical states exhibit counterfactual dependency. Although, with regard to those states that are outside us, we may look to constant conjunction or counterfactual dependence as a sign that they are causally related, the same is not true of the states that we bring about as agents. As Broad (1929) argues with regard to the regularity theory, in such cases 'we can see without waiting for the result that such and such a volition is a necessary condition of such and such a bodily movement' (p.103). Such causal relations do not have to be part of a general pattern for us to form the conclusion that they are causal, and I would suggest that the same consideration applies with regard to counterfactual dependency.

Given that the naïve regularity theory of causation and the counterfactual theory of causation do not provide satisfactory accounts of causation, and, indeed, given that generalist theories of causation in general do not combine well with our acceptance of CPs as the properties of causation, to what theory of causation should we turn? Here I shall sketch two alternative theories of causation. The first is advanced by Ehrling (1997) and analyses causation in terms of CP persistence. The second is proposed by Harré and Madden (1975), and appeals to the notion of non-reductive powers. What should be clear is that neither is biased towards a physicalist interpretation of causation.

2. Ehrling's Theory of Causation

According to Ehrling (1997) 'A causal process is a persisting trope' (p. 122). Ehrling maintains that the causal relata are tropes, or in other words, trope-instantiations (p. 115). The properties of causation are tropes rather than universals because, according to Ehrling, property persistence can only be accounted for by reference to trope persistence. (ch. 4)). Ehrling maintains that there is a causal connection between a cause and its direct effect, if the trope-instantiation that is the effect, is identical with, or partially identical with the trope-instantiation that is the cause. To see what Ehrling means by this, we need to consider the various types of causal process that he considers to exist.

²⁵ Furthermore, the energy transference theory and the generalist theories seem to be trying to achieve different things. The energy transference theory is attempting to advance the empirical identification of causal interactions with energy exchanges. Theories of causation, such as the counterfactual analysis, are attempting to provide a conceptual analysis of causation. I would therefore be reluctant to combine the two.

The most basic causal process involves an individual trope that persists unchanged; 'such unchanging properties form a causal chain linking the past with the future' (p. 14). For example, the rose's redness at t1, is causally connected to the rose's redness at t2, in virtue of the fact that the rose's redness persists from t1 to t2.

More complex causal relations involve patterns of partial trope persistence, and these include trope fission and trope fusion. Because Ehring maintains that there are compound tropes, tropes may display partial persistence. That is, parts of a compound trope may continue to exist, whilst other parts of the compound trope cease to exist – such a compound trope displays a partial trope persistence (p. 14). Partial trope persistence may involve either fission or fusion. 'A fissioning compound trope breaks apart to form simpler tropes, which are partially identical to the original compound trope... Trope fusion is the reverse of trope fission: a compound trope is formed from simpler tropes. The fusing tropes are partially identical to the compound trope' (p. 14).

Ehring maintains that there are causal processes corresponding to trope fission and trope fusion. When a compound trope fissions into separate tropes, we can follow a causal process from the compound trope to the resulting separate tropes. The cause is the compound trope. The effect is the later separate tropes. For example: 'Trope fission is illustrated by the division of a particle into two particles, each with charge less than that of the dividing particles. Here we can trace a causal process from the original charge to the emerging, lesser charges. The charge of the dividing particle is a cause of the later, smaller charges in virtue of the fact that the earlier charge was in part constituted by the charges that emerge in the process' (p. 122).

In the case of trope fusion, separate tropes fuse to become a trope compound. Within fusion-based causal processes we can trace a causal process from the original separate tropes to the resulting compound trope. Hence: 'Trope fusion is illustrated by the collision of two particles that come to form a new particle with a charge equal to the combined charges of the colliding particles. We can trace a causal process from the original charges to the emerging, greater charge. The charges of the colliding particles are causes of the later, greater charge in virtue of the fact that the later charge is in part constituted by the fusion of the charges of the colliding particles' (p. 122). Hence, for Ehring, most causation consists in the building up and breaking down of tropes over time.

Finally, crucial to Ehring's account is the distinction between stable and unstable trope fusions. Whilst a stable trope fusion will remain intact unless disturbed by some outside influence, an unstable fusion will fission without anything operating on it. (p. 119) Hence, two (compound) tropes may fuse to form an unstable complex trope and then fission into either the original fusing tropes or tropes that are constitutive of the fused trope. Consequently, Ehring can allow that some fusion-based causal processes are unstable. In such cases, two or more tropes fuse, to form an unstable compound trope which eventually fissions. In such cases, 'We can trace a causal process back from the emerging fission products to the complex trope and then from that complex trope to the original fusing tropes' (p. 122). To give an example from Ehring (pp.124- 5), Jones cuts a string with a pair of scissors. The cause is the movement and sharpness of the scissors. The effect is two separate-length tropes of string. When the scissors contact the string, the movement/ sharpness trope comes to form part of a fused trope — the trope that is the length of the string fuses with the trope which is the movement/ sharpness of the scissors. This is an unstable fused trope. The unstable fused trope then fissions into the movement/ sharpness of the scissors, and two string half-length tropes.

How should one apply this theory of causation to psychophysical causation? Well, presumably one should say that when an instantiation of a mental CP causes an instantiation of a physical CP, a momentary compound CP is formed consisting of the mental and physical CP. This unstable fusion product quickly fissions, but the physical CP emerges altered. That is, as a result of its fusion with the mental CP, it fissions or fuses in certain ways that it otherwise would not have done.

I do not pretend that this theory of causation is entirely clear or persuasive. In particular, one may question exactly when unstable compound tropes are formed, and when the forming of an unstable compound trope is a causal process. However, my aim is not to offer a brief on behalf of Ehring's account of causation, but show how it differs from the energy transference theory of causation in not being obviously biased towards a physical model of causation. Like the energy transference theory of causation, Ehring maintains that his theory of causation is a reductive singularist theory of causation. It is singularist because causation is primarily a local matter, causation simply consisting in trope persistence (p. 14). It is reductive because causation can be analysed in terms of noncausal facts about trope persistence, fission and fusion.

However, Ehring's account of causation differs from the energy transference theory of causation in a number of significant ways. Although, like the energy transference theory of

causation, Ehring maintains that causation involves a persisting entity that carries causal influence, he disagrees with the energy transference theory about what this persisting entity is. For the energy transference theory, the persisting entity that carries causal influence is energy/ momentum. For Ehring, the persisting entities that carry causal influence are tropes. Thus, like the energy transference theory, Ehring's singularist account of causation provides a local mechanism for the transmission of direct causal influence. But whilst the energy transference theory appeals to the transference of energy/ momentum to explain how causes bring about their direct effects, Ehring appeals to trope persistence. In doing so, Ehring offers a much more general account of the causal mechanism, and importantly one which is not biased towards a physicalist interpretation. In order to explain how a cause brings about an effect one does not need to appeal to the notions of force, energy or momentum; one simply appeals to tropes. Of course, it might be true that all tropes are physical tropes, but Ehring's account of causation does not dictate that this is so.

It is also important to note that, unlike the energy transference theory of causation, for Ehring, causation does not involve any actual transference of this persisting entity. That is, Ehring is not proposing that in order for a cause to bring about an effect, a trope must be transferred from cause to effect. Indeed, if this was Ehring's proposal, then this would prove problematic for those interactive substance dualists who maintain that mental and physical substances share no common properties.²⁶ Rather, causation involves trope persistence. It is true that the trope instantiation that is the effect must be identical with, or partially identical with, the trope-instantiation that is the cause, but because Ehring allows that compound tropes may be unstable, this need not worry such a substance dualism. To see this, compare the compound trope that is the mental substance's instantiation of a mental trope and the body's instantiation of a physical trope, with the compound trope that is the string's instantiation of a length trope and the scissors' instantiation of a movement/ sharpness trope in the above example.

3. *Harré & Madden's Theory of Causation*

Perhaps more plausibly, rather than adopting a singularist theory of causation in which the causal relations reduce to noncausal facts about the causal relata, one could adopt a non-

²⁶ Lowe (2000b, p. 22) raises and rejects this kind of argument as a possible argument against substance dualism.

reductive theory of causation. According to such theories, causation cannot be reduced to noncausal properties and relations of the cause and effect. More specifically, according to the non-reductive singularist there are irreducible causal facts about particulars that cannot be reduced to non-causal facts about particulars or noncausal facts about particulars together with laws.

An example of such a theory is found in Harré and Madden's *Causal Powers*, which argues for the locus of causal powers in the 'powerful particular'.²⁷ According to this theory, causation is to be analysed in terms of powers, but powers do not themselves reduce to noncausal facts about the particulars. To ascribe a power to a thing, is to assert that it can do what it does in virtue of its nature. That is, being of a certain nature endows a thing with the power to manifest itself in certain ways, in appropriate circumstances. (1975, p. 91) However, although the powers of a thing are explained by the nature of a thing, the former should not be eliminated for, or reduced to, the latter (p.112). 'Power' and 'nature' are intimately interwoven and an attempt to assign ontological priorities amongst them is futile. Although powers are ineliminable, they should not be seen as occult or mysterious, because they have an unproblematical basis in the nature of the entities involved. 'The ineliminable, but non-mysterious powers and abilities of particular things, then, are the ontological "ties that bind" causes and effects together...' (p. 11)²⁸

Hence, although this theory appeals to the notion of a power, because it is wholly non-reductive it rejects any attempts to reduce powers to non-causal facts. Obviously, therefore, powers are not to be reduced to non-causal concepts that are associated with the physical domain such as 'energy' or 'force'. Of course, if monism is correct, then it will always be the physical nature of a thing that endows it with powers to act within the

²⁷ Anscombe's (1993) account of causation also falls into the category of being a non-reductive singularist position. With regards to its singularism, partly in virtue of probabilistic causation (p. 101), Anscombe rejects Humeanism, although it is unclear whether she considers that causal relations must be grounded by probabilistic laws. Causality is a matter of 'deriving from' or 'arising of', and as 'deriving from' is nearly synonymous with 'causing' and Anscombe does not provide a reductive account of 'deriving from' this account can be taken as non-reductionist. (See Ehring (1997, pp. 7-8) for further support of this interpretation).

²⁸ Note that this position is compatible with understanding the causal relata as property-instantiations. Harré and Madden themselves maintain that according to their view 'What may be singled out as the cause may be an event, a state of affairs, or even in certain contexts, a material substance.' (1975, p. 5) And given that a substance has the inherent causal powers that it does in virtue of the properties that it instantiates, it seems reasonable to combine this theory of causation with an understanding of the causal relata as property-instantiations. A substance's instantiation of a CP brings with it particular causal powers, hence it is the substance's instantiation of the relevant CP that brings about an effect.

physical domain. But this theory of causation is certainly not biased towards such an interpretation. The mental properties that a substance instantiates may bring with them new causal powers that furthermore make a difference within the physical domain.

Also note that, because of its non-reductive character, Harré and Madden should therefore reject the question of how a cause brings about its direct effect, if by this one is demanding some non-causal explanation of how a cause brings about its direct effects. Beyond the fact that a substance's instantiation of a CP brings with it certain causal powers that enable it to have certain effects, there is no asking how. Consequently, a theory of interactive mentalism which assumes this theory of causation, and which offers an account of psychophysical causation, should not be expected to offer an explanation of how the mental cause brings about this physical effect, if, by this, one is requiring the mentalist to cite a non-causal mechanism.

6.5 Interactive Mentalism

These two theories of causation are, unlike the energy transference theory of causation, not biased towards a physicalist model of causation. With these two theories of causation in mind, I now want to sketch three very different accounts of psychophysical causation. From the perspective of an energy transference theory, these accounts of psychophysical causation look wholly implausible. Indeed, the causal roles that they propose that mental states play within the physical domain would not be recognised by a proponent of the energy transference theory. But from the perspective of the two theories of causation that I have just outlined the same is not true.

Although I would not want to suggest that any of the following theories are problem-free, what does seem to be the case is that it is not obvious that the monist can simply respond to them by rehearsing the argument from causal overdetermination. This is because it is not obvious that physics can appeal to physical causes in order to fill the 'causal gaps' that these theories of interactive mentalism suggest exist within the physical domain. Mental causes seem better suited to play these causal roles.

I present the theories of interactive mentalism in the order that I do, because they are consistent with increasingly stronger causal closure principles. Hence, in order for the monist to rule them out by way of the argument from causal overdetermination, they are

required to support increasingly stronger, and hence less plausible, understandings of the causal closure principle.

1. The Collapse of the wave-function

I start with the most contentious proposal. This is London and Bauer (1983), Wigner's (1961), and more recently Stapp's (1997) proposal, that consciousness causes the collapse of the wave function. The monist's appeal to current physics, in order to support the causal closure of the physical domain, would suggest that within all, or at least the best, interpretations of the formalism of quantum mechanics the causal closure principle functions as a working premise, and certainly that none appeals to mental properties as causes of physical a effects. However, some interpretations of quantum mechanics actually require that a place for mental causes is given.

In order to be at all plausible, quantum mechanics must be interpreted in such a way that it is possible to reconcile the classical behaviour and appearance of everyday physical objects that we experience, with the world of superposed wave functions described by the formalism of quantum mechanics. The group of interpretations of the formalism of quantum mechanics that are of particular interest to the mental causation debate are those that maintain that superposition at the micro-level leads us to experience unsuperposed discrete structures in the macro-world due to a collapse in the wave function. If a measurement of a micro-system is made, which is suitable for distinguishing between results that are possible but not certain, only one of the possible results will be measured. However, until the act of measurement, all of the results remain possible. The suggestion is that the existence of a range of possibilities before the system is measured, is not due to a lack of knowledge, and that after measurement, the fact that there is a definite result is an objective feature of the microsystem. When a measurement has been made, the wave function collapses into a state with a definite value. This explains why, when a micro-system is measured, a definite result is gained.

If it is an act of measurement that causes the wave function to collapse, what this suggests for the causal closure principle depends upon one's analysis of measurement. That is, on what it is about measurement, as opposed to any other interaction, that causes collapse. Arguably, what is special about a measurement interaction is that a measurer must observe and thus become aware of the values determined by the measuring process. This has led London and Bauer, Wigner and Stapp, to maintain that it is the act of observation that causes the wave function to collapse. When the result of a

measurement is observed by a consciousness, or in other words, when a quantum state affects someone's consciousness, this causes the wave function to collapse.²⁹

What it is about consciousness that causes collapse? That is, what property of consciousness makes this causal difference? One obvious candidate is the property of intentionality. Thus, for example, it is the beliefs that the observer forms in observing the measuring apparatus that causes the wave function to collapse. On the other hand, Stapp indicates that it is the experiencing of the value of a measurement, the phenomenal aspect of consciousness, that causes collapse.³⁰ And London and Bauer, maintain that it is the property of introspection, the ability of the observer to attend to himself in abstraction from the physical system with which he interacts, that intimately links consciousness with collapse. These suggestions are at best tentative.

It is important to note, that this theory should not be taken as a disguised form of monism. If a mental state causes the collapse of the wave-function, then this mental state must be attributed causal powers that exist over and above the causal powers of physical a properties. That is, it must be understood as an emergent cause. Hence, in analysing the role of the observer in the collapse of the wave function, one should not trace the process of observation ever deeper into the physiology of the observer if this leads one to see the observer as just another macrosystem, whose causal efficacy is nothing over and above the causal efficacy of the quantum system from which it is ultimately composed. This is because there is nothing within quantum mechanics that one can single out as a cause of collapse. According to this theory, contrary to the monist, consciousness is not something that is to be explained away ultimately in terms of quantum mechanical processes, but is taken as a given, its properties making a causal difference within quantum mechanical processes. Consequently, given such a position, physics is not causally closed in anything but a trivial sense.

Note also that the claim is not that physics needs to appeal to mental properties in order to provide a complete causal account of paradigm physical effects. Rather, the claim is that physics needs to appeal to mental properties in order to provide a complete causal account of physical a effects. This consideration leads Papineau to comment that 'It would

²⁹ To quote Stapp; 'Chalmers suggests that perhaps there is a small loop-hole in quantum theory that might provide an opening for consciousness. But there is not just a small loop-hole: there is a gigantic gap, which consists of *fully half of the theory*, and this hole provides an ideal home for consciousness' (1997, p. 212).

³⁰ Stapp — 'This makes the experiential aspect of the actualization events the *cause* of the classical character of the collapse events...' (1997, p208).

consider the shift from universals to particulars to allow one to advance a new form of psychophysical reductionism that does not encounter the problem of multiple realisability. This is because I do not consider that one can plausibly combine a trope monism with a type dualism in the way that they require. However, I argued that this does not mean that the shift can be ignored. What the properties of causation are has a number of general but very important effects on the mental causation debate. In the first place, what it is for a property to exist, and for one property to be identical with another, as well as the question of what dependence relationships exist between properties, all affect the mental causation debate. But in order to provide an analysis of properties, clearly one must first decide upon what properties are. Secondly, one's theory of causation is heavily influenced by whether one understands the properties of causation to be universals or characterising particulars. In particular, it is implausible to combine an understanding of the properties of causation as characterising particulars with a generalist theory of causation. If the properties of causation are characterising particulars, one is thus limited to either a reductive or a non-reductive singularist position. This has a knock on effect within the mental causation debate, because one's understanding of psychophysical causation is constrained by the theory of causation that one is basing psychophysical causation on.

The second part of this thesis was concerned with the first of these issues, that of providing an analysis of characterising particulars. I argued that for one characterising particular to be identical with another, they must belong to the same exact resemblance class and be instantiated by the same substance. To the further question of when one characterising particular exactly resembles another, I rejected any semantical approach, instead suggesting that they must have exactly resembling causal powers. Similarly, a causal criterion of characterising particulars' existence was thought to be far more plausible than a semantical one.

The rejection of a semantical analysis of properties has a number of important affects within the mental causation debate. In particular, it leads one to question the multiple realisability argument — it is arguably mental predicates, not mental types, that are multiple realised. It also allows one to reject Kim's Principle of Explanatory Exclusion. More generally, it leads one to recognise that from the fact that there are different levels of predicates, one should not conclude that there are different ontological levels of

substance dualism the only kind of interactive mentalism.³¹ And to reject all forms of interactive mentalism out of hand is merely suggestive of a physicalist prejudice.

Another reason why this proposal is viewed with suspicion is that it seems causally mysterious. In what sense of 'cause' is the collapse in the wave function caused? How could a mental state cause a collapse in the wave function? What is the mechanism that is involved in such a causal relation?

The idea that some state (regardless of whether it is mental) 'causes' the collapse of the wave function, will appear unintelligible to those who assume an energy transference theory of causation. According to this theory, for one state to cause another, there must be an energy transfer from the cause to the effect. But the having of a particular mental state and the resulting collapse in the wave function is one particular case that cannot involve an energy transfer. Although energy is conserved in the Schrödinger evolution of quantum systems, it is violated by wave collapses. Consequently, here one clearly cannot analyse causation in terms of energy transfer. Hence, for those who assume an energy transference theory of causation in all other cases of physical causation, the proposal that consciousness causes the collapse of the wave function will seem mysterious and strange.

If, however, one does not assume an energy transference theory of causation, the same is not true. Causation, qua causation, does not involve energy transfer, and hence these theories of causation are wholly consistent with the possibility that causal relations are not always accompanied by an energy transfer. For example, take Harré and Madden's theory of causation. Given this theory of causation, one could allow that in virtue of its nature, consciousness has the causal power to cause the collapse of the wave function. This is a brute power of consciousness. Consciousness causes the collapse of the wave function by manifesting this causal power. Any further question about how consciousness causes the collapse of the wave function, that is, about what mechanism is involved, should be rejected, for according to this non-reductive analysis of causation there is no such noncausal process linking a cause with its direct effect.

³¹ Stapp's, for example, maintains an interactive property mentalism. He considers consciousness to be the cause of the collapse of the wave function, where consciousness is a high level process '...rising out of the matter-like aspects of nature lies another dynamics governed by the experiential aspects of nature.' (1997, p.213).

Admittedly, to reject an energy transference theory of causation does not leave the claim that consciousness causes the collapse of the wave function wholly intelligible. But this is not due to the fact that no transfer of energy is involved or that no non-causal mechanism can be given for such a causal relation. It is instead due to the fact that it is hard to make sense of the causal chain of states involved. It is wildly counterintuitive that the effect on consciousness of the quantum mechanical system is direct. Rather it will happen through the intermediate apparatus of the sense organs, nerves and brain of the observer's body. Obviously this process is not spontaneous, and yet, when one observes the contents of Schrödinger's box, one does not experience an initial state of superposition (indeed, would it be possible for one to experience such a thing?), or an initial state of vagueness whilst the physical input is in a superposed state. This should perhaps lead us to have reservations about this proposed form of psychophysical causation.

I shall now turn to two far more plausible forms of interactive mentalism. Unlike the above proposal which suggests that mental causes have physical a effects, according to these remaining proposals mental causes have paradigm physical effects.

2. Broad's theory of psychophysical interactionism

Broad (1929, p. 104) considers the laws of the conservation of energy to be 'irrelevant' to the question of whether interactive mentalism is correct. In response to those who believe the conservation laws of physics to threaten psychophysical interactionism, he suggests that mental states could 'determine that at a given moment so much energy shall change from the chemical form to the form of bodily movement; and they determine this, so far as we can see, without altering the total amount of energy in the physical world' (p. 109).³² Broad's claim seems to be that mental states prompt, or if you prefer conduct, the transfer of energy between physical states. However, they do not do so by transferring energy. That is, they do not prompt energy transfers by transferring energy. (Indeed, if they did prompt energy transfers by transferring energy, one might justifiably ask what prompted the energy transfer that prompted the energy transfer).

Broad's proposal suggests that although there may be no gaps in the physical chains of energy transfer, there is an additional causal role to be played by mental states within the physical system. This is a causal role that would obviously not be recognised by the

³² Although, note that the second law of thermodynamics might be thought to threaten Broad's proposal.

energy transference theory of causation. However, this theory of psychophysical causation is certainly compatible with other theories of causation.

It is, for example, perfectly consistent with Harré and Madden's account of causation that, in virtue of their natures, mental particulars have the brute power to prompt transfers of energy between physical states. There is, I am sure, no need to repeat the fact that to raise the question of how the mental state prompts the transfer of energy, would be deemed an illegitimate one within this theory of causation.

Turning to Ehring's analysis of causation in terms of trope persistence, if one tries to embed Broad's proposed form of psychophysical causation into Ehring's account of causation, matters are a little more complicated, largely because of the complicated nature of Ehring's account of causation. Let us say that a mental state *M*, prompts the transfer of energy from neurological state *N1* to neurological state *N2*. Neurological state *N1* is the instantiation of a complex trope consisting of quantity of energy *x*. Neurological state *N2* is the instantiation of a complex trope consisting of quantity of energy *z*. Quantity of energy *x – y* is transferred from *N1* to *N2*. Given Ehring's account, the mental trope that characterises mental state *M* fuses with the complex trope consisting of quantity of energy *x*, to form an unstable compound trope. This compound trope promptly fissions into the mental trope, a trope quantity of energy *x – y*, and a trope consisting of quantity of energy *y*. Upon this fission, the trope that is the quantity of energy *x-y*, fuses with the trope that *N2* instantiates, forming a compound trope consisting of a trope that is the quantity of energy *x- y* and a trope that is the quantity of energy *z*. Note that Ehring's account allows the question of how a mental cause prompts the transfer of energy. The answer is by trope fusion.

Hence, unlike the energy transference theory of causation, these theories of causation do not provide a barrier to Broad's proposed form of psychophysical interactionism. Broad's proposal is clearly intelligible given these theories of causation. Now, of course, this is not to suggest that Broad's theory does not face problems. In particular, one might plausibly ask what need there is for mental states to play this causal role within physical systems. Do all physical systems (not just the neurological system) require that their energy transfers are prompted? If so, then unless we are to admit panpsychism, in these other physical systems it is presumably physical states that prompt the energy transfers. But then if physical states can prompt energy transfers in these other systems, why can't they prompt energy transfers in neurological systems — that is, why do we need to appeal to mental states to play this causal role in neurological systems?

What is needed is some reason to suggest that transfers of energy are only prompted within neurological systems. (I am not suggesting that this is Broad's position). But, of course, if this is the case, the question is why? What differentiates neurological systems from all other physical systems? One possible response is to appeal to the idea that in neurological systems, certain chains of transfers of energy are directed towards a goal. I desire that I should raise my arm, and my arm goes up. Any causal chain involving this desire would seem to be directed at the end result of raising my arm. In other systems, chains of causation seem to lack this directionality. This directionality could then be explained by appealing to the idea that in the former case, the energy transfers are prompted. Why can not physical states within neurological systems play this causal role? Why do we need to appeal to mental states to fill this 'causal gap' within the physical system? Well, the obvious reply is that mental states such as beliefs and desires are intentional states and it is the intentionality of a mental state, the mental's directedness on its objects, that allows it to play this additional causal role of prompting energy transfers.

Clearly, this is only meant to offer a sketch of how one might develop the claim that mental causes prompt energy transfers, and nor am I suggesting that it is problem-free. But what I think this does show is that, given the rejection of an energy transference theory of causation, one can recognise the possibility of there being additional causal roles within the physical domain, which mental states might be better suited to play than physical states.

3. Lowe's theory of psychophysical interactionism

I now wish to turn to Lowe's proposal (1996, 1999a, 2000), which was discussed in Part One, §6. It was observed that Lowe does not consider mental events to cause bodily behaviour by initiating a particular chain of physical events, because he argues that there is nowhere that one could place the mental cause within a causal chain of physical events, such that it would be correct to say that this mental event initiates this particular bodily behaviour. Lowe would therefore presumably reject Broad's proposal that mental states prompt energy transfers, because this proposal would seem to imply that mental events cause bodily behaviour by acting upon specific points within the intermeshed causal chains between the mental event and the bodily behaviour.

According to Lowe, a mental event causes the particular pattern of events, ending in a piece of bodily behaviour, to exist. Importantly, it does not do this by causing a particular

physical event and thereby initiating a sequence of physical events ending in the bodily behaviour. It will be remembered that to capture this causal role of the mental, Lowe distinguished between event causation and fact causation. A mental event is necessary and sufficient for the fact that this particular pattern of events terminating in the bodily behaviour exists.

Lowe maintains that this additional causal role is required, because otherwise, the fact that this causal tree of physical events converges upon a particular event — the bodily behaviour — looks purely coincidental. Mental events render this non-coincidental. A mental event such as a belief or desire is able to play this causal role, precisely because it has as its intentional object the relevant bodily behaviour. The mental event is directed upon the occurrence of this particular bodily behaviour. (In particular, see Lowe (1999a, p. 234-9). Consequently, in response to the monistic objection that even if some event was required to play this causal role within the physical domain, one would never need to appeal to mental events to do so, Lowe is able to respond that it is the intentional nature of mental events that enable them to play this causal role.

In Part One §6, I rejected Lowe's theory because I argued that one could not plausibly deny the homogeneity of the causal relata. Here I wish to consider whether one can capture the central ideas of Lowe's account, despite accepting the heterogeneity of the causal relata.

Clearly, Lowe does need to say what theory of causation he is assuming in his discussion of psychophysical causation, and furthermore whether event-event (physical) causation is the same kind of causation as event-fact (psychophysical) causation. It is first important to note that Lowe's theory of psychophysical causation is inconsistent with an energy transference theory of causation. In the first place, even if one denies the homogeneity of the causal relata, if both event-event causation and event-fact causation are based within an energy transference theory of causation then one still encounters problems with regard to the conservation laws of physics. More importantly, that Lowe would clearly not want to maintain an energy transference of causation, at least in the case of event-fact causation (or, in other words, psychophysical causation), is suggested by the way that Lowe considers mental causes to make a difference in the physical domain. An energy transference theory requires that energy is transferred from the cause to a specific event, or that at the very least that it disperses its energy to a specific set of events. But that mental events cause bodily behaviour by acting upon a specific physical event in the causal chain leading up to this bodily behaviour, or indeed that they act upon a particular

set of physical events within this causal web, is precisely what Lowe is denying. Hence, within Lowe's account there is no specific physical event or set of physical events that a mental event could be said to transfer energy to.

More generally, Lowe's account of psychophysical causation is incompatible with any transference theory of causation, for all share the assumption that there is a persisting entity that is transferred to a specific entity or is dispersed between a set of entities. Similar considerations must also lead one to conclude that Lowe's account of psychophysical causation is incompatible with Ehring's theory of causation, for according to Lowe, there is no physical particular or set of physical particulars that a mental particular could be correctly said to fuse with.

However, not all theories of causation require that there is a specific event, or set of events, upon which a cause acts. Hence, consider Harré and Madden's theory of causation. Given this theory of causation, one could allow that mental events have the causal power to make a causal tree of physical events converge upon a particular physical event, and that mental events have this special causal power in virtue of their intentional nature. Given this theory of causation, it is not further required that mental causes exert this causal power upon a particular physical event within this causal chain of physical events. To the comment that this kind of causation seems strange, one can respond that this is because most causation is not like this. In normal, everyday physical causation one can always identify a specific event or set of events upon which a cause acts, arguably because such causation always involves transfers of energy. But given Harré and Madden's theory of causation, it is not the case that all causation must be like this. Psychophysical causation is different in nature from such purely physical causation, and mental states are able to play this special causal role precisely because of their intentional nature. Now if this does adequately capture Lowe's picture of psychophysical causation, then Lowe's distinction between event and fact causation seems to become redundant. If, on the other hand, this is not a satisfactory account of psychophysical causation, then Lowe owes us an account of what kind of causation event-fact causation could possibly be.

To conclude, once one recognises that the energy transference theory of causation should not be applied to psychophysical interactions, one releases the mentalist from a model of causation that has served to hinder his progression towards a plausible theory of psychophysical causal interaction. One will be in a position to develop a more subtle understanding of the way in which mental states may be causally efficacious within the

physical world. The kinds of psychophysical interactionism proposed by Wigner et al, by Broad, and by Lowe, provide three ways in which the interactive mentalist might attempt to do this. Although these three proposals greatly differ, I hope to have shown that all of them are inconsistent with an energy transference theory of causation, and yet can be made plausible, at least from a causal perspective, by basing them within a less physically biased theory of causation. And although none of these theories of interactive mentalism are without their problems, what is clear is that, unlike Hart's theory of interactive mentalism, it is not at all obvious that the monist can simply respond to these forms of interactive mentalism by rehearsing the argument from causal overdetermination. This is because it is far from evident that physical states could fill the specific causal gaps within the physical domain that these theories of interactive mentalism propose there to be.

Summary

Within this discussion, I have argued that, contrary to indications, the mental causation debate cannot detach itself from metaphysics, because metaphysics provides the very framework on which to base one's theory of mental causation. Neither will it do to make ad hoc metaphysical assumptions, for if the ontological framework upon which one is basing one's theory of mental causation is unsound, one's theory of mental causation will inevitably prove to be unsatisfactory.

With an independently motivated, plausible ontological system, one has an invaluable tool within the mental causation debate. It enables one to reject many of the positions in the mental causation debate that (implicitly) depend upon unsound ontological assumptions, to dismiss many of the problems that have typically worried those within the mental causation debate, and to recognise what is truly at issue amongst those theories that remain.

I have argued that there are three broad metaphysical issues that are pivotal to the mental causation debate. Firstly, what causation is a relation between. Secondly, what a plausible analysis of properties should consist in, which includes consideration of what it is for a property to exist, what it is for one property to be identical with another, and consideration of the various other relations that properties can enter into with each other. Thirdly, what the causal relation is. That is, the question of how to understand causation itself.

With regard to the first issue, that of the causal relata, I have suggested that there are properties, and furthermore that these play an essential role within causation. More contentiously, I have argued that the causal relata are most plausibly property-instantiations, where properties are here to be understood as characterising particulars.

This affects the mental causation debate in a number of important ways. Most obviously, it leads to the rejection of Davidson's anomalous monism, because the plausibility of this theory of mental causation is wholly dependent upon Davidson's theory of the causal

relata, which is developed from within a nominalist ontology. Davidson's theory of the causal relata has rarely been properly respected within the mental causation debate. Philosophers of mind have mistakenly accused Davidson's anomalous monism of leading to some kind of property epiphenomenalism. Contrary to this, Davidson's anomalous monism does not lead to property epiphenomenalism, regardless of whether the term 'property' is here being interpreted in an ontological or a linguistic sense. Davidson's theory of mental causation is unsound because of the implausible ontological system within which it is based.

Whilst Davidson's theory of mental causation fails because it assumes an implausible theory of the causal relata, I have suggested that other theories of mental causation are unsatisfying because they rest upon an understanding of the causal relata that is vague and unclear. This is the case with Macdonald's non-reductive physicalist response to the problem of mental causation. It is far from evident what understanding of the causal relata Macdonald is presupposing, and once one tries to unpack Macdonald's ontological proposal, it becomes apparent that it cannot provide any real advancement within the mental causation debate.

The other position within the mental causation debate that I have suggested makes questionable claims about the causal relata, is Lowe's interactive mentalism. Lowe's arguments for interactive mentalism have much to recommend them. However, Lowe's response to the problem of causal overdetermination, namely that of denying the homogeneity of the causal relata, is, I have argued, very difficult to support from an ontological point of view. That is, it is implausible that one can advance an understanding of an 'event' and a 'fact' that will enable one to motivate the distinction between event causation and fact causation that Lowe requires. And obviously, if there is no such way of doing this, Lowe's solution to the problem of mental causation must be rejected.

Certainly, my most contentious ontological claim is that the properties that the causal relata instantiate are particulars rather than universals. That properties are universals, is an (implicit) assumption that is common to most within the mental causation debate. Indeed, as Heil and Robb (forthcoming) observe, it has a status within the philosophy of mind such that it deserves to be called a dogma. Contrary to Heil and Robb, I do not

consider the shift from universals to particulars to allow one to advance a new form of psychophysical reductionism that does not encounter the problem of multiple realisability. This is because I do not consider that one can plausibly combine a trope monism with a type dualism in the way that they require. However, I argued that this does not mean that the shift can be ignored. What the properties of causation are has a number of general but very important effects on the mental causation debate. In the first place, what it is for a property to exist, and for one property to be identical with another, as well as the question of what dependence relationships exist between properties, all affect the mental causation debate. But in order to provide an analysis of properties, clearly one must first decide upon what properties are. Secondly, one's theory of causation is heavily influenced by whether one understands the properties of causation to be universals or characterising particulars. In particular, it is implausible to combine an understanding of the properties of causation as characterising particulars with a generalist theory of causation. If the properties of causation are characterising particulars, one is thus limited to either a reductive or a non-reductive singularist position. This has a knock on effect within the mental causation debate, because one's understanding of psychophysical causation is constrained by the theory of causation that one is basing psychophysical causation on.

The second part of this thesis was concerned with the first of these issues, that of providing an analysis of characterising particulars. I argued that for one characterising particular to be identical with another, they must belong to the same exact resemblance class and be instantiated by the same substance. To the further question of when one characterising particular exactly resembles another, I rejected any semantical approach, instead suggesting that they must have exactly resembling causal powers. Similarly, a causal criterion of characterising particulars' existence was thought to be far more plausible than a semantical one.

The rejection of a semantical analysis of properties has a number of important affects within the mental causation debate. In particular, it leads one to question the multiple realisability argument — it is arguably mental predicates, not mental types, that are multiple realised. It also allows one to reject Kim's Principle of Explanatory Exclusion. More generally, it leads one to recognise that from the fact that there are different levels of predicates, one should not conclude that there are different ontological levels of

properties. This consideration is, I suggested, of particular importance in assessing the plausibility of non-reductive physicalism.

Indeed, one of the central concerns within the second part of my thesis was that of non-reductive physicalism. Distinguishing between a non-reductive physicalism that considers mental and physical properties to be instantiated by the same object, and one that considers them to be instantiated by different objects (a whole and its parts), our discussion began with a consideration of the former kind of position. Our first concern was with establishing whether this kind of non-reductive physicalism really could establish a dependence relation between mental and physical properties by appealing to supervenience. Starting with the now familiar consideration that co-instantive supervenience lacks the directionality of a dependence relationship, I considered the ways in which ontological dependency is commonly formulated so that it captures an asymmetry, and argued that psychophysical supervenience claims cannot be strengthened in a similar way. I then went on to consider the argument that although co-instantive supervenience is not a dependence relationship, a dependence relationship must be appealed to in order to explain it, and hence the non-reductive physicalist's claim that mental properties are determinable properties or, alternatively, second-order properties.

However, the rejection of a semantical analysis of properties led me to question whether the claim that the mental and physical are related, for example, as determinable to determinate, really allows one to advance a non-reductive physicalism. Heil and Robb's rejection of a semantical criterion leads them to reject all property layers within an object, hence, I suggested, leading to the rejection of all forms of non-reductive physicalism that appeal to a co-instantive supervenience account. Contrary to Heil and Robb, I argued that to reject the claim that there are property-levels within an object, it is not enough to reject a semantical analysis of properties. To come to any conclusions about this matter, one needs a positive analysis of CPs.

And turning to a causal analysis of CPs, I argued that this in fact brings into question all forms of non-reductive physicalism, due to the combination of premises that the non-reductive physicalist is committed to. Furthermore, with the rejection of a semantical analysis of properties, one should be suspicious of the non-reductive physicalists' attempts to secure the causal efficacy of the mental by denying the homogeneity of

causation. Taking Yablo's non-reductive physicalism as my central example, I argued that all that Yablo's account really points to is that mental predicates play a non-redundant role within causal explanations. Indeed, if the mental and physical are related as determinable to determinate, this is, given a plausible analysis of properties, to actually advance an eliminativism. These considerations about Yablo's account indicate that we should be suspicious of any non-reductive physicalism that attempts to avoid epiphenomenalism by denying the homogeneity of causation.

This part of my thesis ended with a discussion of a different property layering — one between different objects. After considering in what way one property could plausibly be said to constitute another, I argued that this notion of property dependence was in fact consistent with three different positions within the mental causation debate; non-reductive physicalism, psychophysical reductionism and an interactive property mentalism. I argued that our causal analysis of CPs leads us to reject the first option, and whether one advances a psychophysical reductionism or an interactive property mentalism, wholly depends upon whether one thinks that there is downwards causation. Hence, property analysis reveals that the true contenders within the mental causation debate are interactive mentalism and psychophysical reductionism. Which of these positions is correct depends upon whether a causal closure principle that is strong enough to establish psychophysical reductionism is plausible.

Part Three focused upon the plausibility of such a causal closure principle. I began by distinguishing between two different strengths of causal closure principle. Within strong causal closure arguments, the causal closure principle is strong enough to allow one to conclude that mental causes are identical with physical causes, merely if combined with the further premise of psychophysical causation. Weaker causal closure principles are of the type appealed to within the argument from causal overdetermination. In response to Lowe's objection that weak causal closure principles are in fact consistent with interactive mentalism, the standard formulations of these weak causal closure principles were modified to capture the strength of argument required by the monist.

The plausibility and strength of a causal closure principle is affected by what one understands a physical cause to be, and hence how one defines the term 'physical.' But it is also affected by what one understands a 'cause' to be, and hence the theory of

causation in which one embeds the mental causation debate. This is not only true of strong causal closure principles, but also weak causal closure principles, for what it is to be a 'cause' and what it is to be an 'effect', and thus what it is to be a 'sufficient cause of an effect', depends upon one's theory of causation. The stronger one's formulation of the causal closure principle, and the less physically biased one's understanding of causation, the harder it will be to provide an understanding of the term 'physical' which is both non-trivial and plausibly causally closed.

After considering how one might plausibly define the term 'physical' for the purpose of the mental causation debate, I suggested that any plausible argument for any type of causal closure principle must be based upon empirical considerations, and that how plausible these considerations are, partly depend upon the theory of causation that is being assumed. In particular, I argued that the monist's (implicit) acceptance of an energy transference theory of causation lies behind many of the empirical arguments for all strengths of causal closure principle. Contrary to their suggestions, monists often make assumptions about causation that are most compatible with an energy transference theory of causation. At worst, the energy transference theory of causation leads to the adoption of the question-begging, strong causal closure claims. But at the very least, in maintaining an energy transference theory of causation, one makes the problem of mental causation look like a very hard problem for the interactive mentalist to resolve. This is because such a theory of causation arises from initial considerations of a model involving purely physical interactions. Hence one is attempting to make sense of psychophysical causation on the basis of a model of causation designed to explain purely physical causation. It only seems highly plausible that the mental will never be needed to account for physical effects, because the energy transference theory of causation is being implicitly appealed to, to determine what a 'causal gap' in the physical domain must be like and how it must be filled. If one bases one's theory of psychophysical causation in a less physically biased theory of causation, then even the weakest kind of causal closure principle that would support monism, loses much of its appeal.

Bibliography

- Achinstein, P. 1983: *The Nature of Explanation*. Oxford: Oxford University Press.
- Alexander, H. (ed.) 1956: *The Leibniz-Clarke Correspondence*. Manchester: Manchester University Press.
- Alexander, S. 1920: *Space, Time, and Deity*. Vol.2. London: Macmillan.
- Anscombe, G. E. M. 1993: 'Causality and Determinism,' in Sosa, E. and Tooley, M. (1993).
- Armstrong, D. M. 1972: 'Materialism, Properties, and Predicates'. *Monist*, 56, pp. 163-76.
- 1978: *Universals and Scientific Realism*, 2 vols. Cambridge: Cambridge University Press.
- 1983: *What is a Law of Nature?* Cambridge: Cambridge University Press.
- 1989: *Universals: An Opinionated Introduction*. Boulder, CO: Westview Press.
- 1997: *A world of states of affairs*. Cambridge: Cambridge University Press.
- Armstrong, D. M., Martin, C. B. and Place, U. T. Crane, T. (ed.) 1996: *Dispositions: A Debate*. London: Routledge.
- Aronson, J. 1971: 'On the Grammar of "Cause"'. *Synthese*, 22, pp. 414-430.
- Baker, G. & Morris, K.J. 1996: *Descartes' Dualism*. London: Routledge.
- Baker, L. R. 1993: 'Metaphysics and Mental Causation', in Heil & Mele (1993).
- Beauchamp, T. L. & Rosenberg, A. 1981: *Hume and the Problem of Causation*. New York: Oxford University Press.
- Beckermann, A. et al. (eds.) 1992: *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Berlin: De Gruyter.
- Bennett, J. 1988: *Events and Their Names*. Oxford: Clarendon Press.
- Bickle, J. 1996: 'New Wave Psychophysical Reductionism and the Methodological Caveats'. *Philosophy and Phenomenological Research*, vol. LVI, No. 1, pp. 57-78.
- Bilodeau, D. J. 1997: 'Physics, Machines and the Hard Problem', in Chalmers (1997).
- Block, N. (ed.) 1980: *Readings in the Philosophy of Psychology*, 2 vols. London: Methuen.
- Broad, C. D. 1929: *The Mind and its Place in Nature*. London: Routledge & Kegan Paul.
- Burge, T. 1993: 'Mind-Body Causation and Explanatory Practice', in Heil & Mele (1993).
- Campbell, K. 1991: *Abstract Particulars*. Oxford: Basil Blackwell.

- Carnap, R. 1947: *Meaning and Necessity*. Chicago: University of Chicago Press.
- Carroll, J. 1994: *Laws of Nature*. Cambridge: Cambridge University Press.
- Castaneda, H. 1984: 'Causes, Causity and Energy'. *Midwest Studies in Philosophy*, 9, pp. 17-27.
- Chalmers, D. 1996: *The Conscious Mind*. Oxford: Oxford University Press.
- Chalmers, D. (ed.) 1997: *Explaining Consciousness*. Cambridge: MIT Press.
- Chisholm, R. M. 1970: 'Events and Propositions'. *Nous*, 4, pp. 15-24.
- Clatterbaugh, K. 1999: *The Causation Debate in Modern Philosophy 1637-1739*. London: Routledge.
- Crane, T. 1991a: 'Why Indeed? Papineau on Supervenience'. *Analysis*, 51, pp. 32-37.
- 1991b: 'All God Has To Do'. *Analysis*, 51, pp. 235-244.
- 1992: 'Mental Causation and Mental Reality'. *Proceedings of the Aristotelian Society*, 92, pp. 185-202.
- 1995: 'Mental Causation'. *Proceedings of the Aristotelian Society*, Supp. Vol. 69, pp. 211-36.
- 2001: *Elements of Mind*. Oxford: Oxford University Press.
- Crane, T., & Mellor D. H. 1990: 'There is No Question of Physicalism'. *Mind*, 99, pp. 185-206.
- Davidson, D. 1967: 'The Logical Form of Action Sentences', in , N. Rescher (ed.), *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh.
- 1967b: 'Truth and Meaning'. *Synthese*, 17, pp. 304-23.
- 1980a: *Essays on Actions and Events*. Oxford: Clarendon Press.
- 1980b: 'Actions, Reasons and Causes', in Davidson (1980a).
- 1980c: 'Causal Relations', in Davidson (1980a).
- 1980d: 'Mental Events', in Davidson (1980a).
- 1980e: 'Psychology as Philosophy', in Davidson (1980a)
- 1980f: 'The Material Mind', in Davidson (1980a)
- 1984a: *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- 1984b: 'On the Very Idea of a Conceptual Scheme', in Davidson, D. (1984a).
- 1984c: ' True to the Facts', in Davidson, D. (1984a).
- 1993: 'Thinking Causes', in Heil & Mele (1993).
- Dretske, F. 1977: 'Referring to Events', *Midwest Studies in Philosophy*, 2, pp. 369-78.
- 1993: 'Mental Events as Structuring Causes of Behaviour', in Heil & Mele (1993).
- Ducasse, C. J. 1940: 'Propositions, Opinions, Sentences, and Facts'. *Journal of*

Philosophy, 33, pp.701-11.

1968: *Truth, Knowledge, and Causation*. London: Routledge & Kegan Paul.

Earman, J. 1986: *A Primer on Determinism*. Dordrecht: D. Reidel Publishing Company.

Eccles, J. C. & Popper, K. R. 1977: *The Self and Its Brain*. Berlin & New York: Springer International.

Ehring, D. 1997: *Causation and Persistence: A Theory of Causation*. Oxford: Oxford University Press.

Fair, D. 1979: 'Causation and the Flow of Energy'. *Erkenntnis*, 14, pp. 219-250.

Fales, E. 1990: *Causation and Universals*. London and New York: Routledge.

Fodor, J. 1993: 'Making Mind Matter More,' in Heil and Mele (eds.).

1994: 'The mind-body problem', in Warner and Szubka (eds.).

Foster, J. 1991: *The Immaterial Self*. London: Routledge.

Goldman, A. I. 1970: *A Theory of Human Action*. Englewood Cliffs, N.J.: Prentice-Hall.

Hahn, L. E. (ed.) 1997: *The Philosophy of Roderick M. Chisholm*. Peru Ill.: Open Court.

Harré, R. & Madden, E. H. 1975: *Causal Powers: A Theory of Natural Necessity*. Oxford: Basil Blackwell.

Hart, W. D. 1988: *The Engines of the Soul*. Cambridge: Cambridge University Press.

Hausman, D. M. 1998: *Causal Asymmetries*. Cambridge: Cambridge University Press.

Heil, J. 1992: *The Nature of True Minds*. Cambridge: Cambridge University Press.

1998: 'Supervenience Deconstructed'. *European Journal of Philosophy*, 6, pp. 146-55.

1998b: *Philosophy of Mind*. London: Routledge.

Heil, J., & Mele, A. R. (eds.) 1993: *Mental Causation*. Oxford: Clarendon Press.

Heil, J. & Martin, C. B. 1999: 'The Ontological Turn'. *Midwest Studies in Philosophy*, XXIII, pp. 34-60.

Heil, J & Robb, D. Forthcoming: 'Ontology and Mental Properties'.

Hempel, C. G. 1980: 'Comments on Goodman's *Ways of Worldmaking*'. *Synthese*, XLV, pp. 193-99.

Hendry, R. F. 1999: 'Molecular Models and the Question of Physicalism'. *An International Journal for the Philosophy of Chemistry*, 5, pp. 117-134.

Hodgson, D. 1991: *The Mind Matters: Consciousness and Choice in a Quantum World*. Oxford: Clarendon Press.

Honderich, T. 1982: 'The Argument for Anomalous Monism'. *Analysis*, 42, pp. 59-64.

1983: 'Anomalous Monism: Reply to Smith'. *Analysis*, 43, pp. 147-149.

- 1988: *A Theory of Determinism: The Mind, Neuroscience, and Life-Hopes*. Oxford: Clarendon Press.
- Horgan, T. 1982: 'Supervenience and Microphysics'. *Pacific Philosophical Quarterly*, 63, pp. 29-43.
- 1993: 'From Supervenience to Superdupervenience.' *Mind*, 102, pp. 555-586.
- Hume, D. A. 1738: *A Treatise of Human Nature*. Rpt. Oxford: Clarendon Press, 1966.
- 1748: *An Inquiry Concerning Human Understanding*. Rpt. Indianapolis: Bobbs-Merrill, 1955.
- Humphreys, P. 1996: 'Aspects of Emergence'. *Philosophical Topics*, 24, pp. 53-70.
- Jackson, F. 1995: 'Mental Properties, Essentialism and Causation'. *Proceedings of the Aristotelian Society*, 95, pp. 253-68.
- 1996: 'Mental Causation.' *Mind*, 105, pp. 377-413.
- 1998: *From Metaphysics to Ethics*. Oxford: Oxford University Press.
- Kim, J. 1990: 'Explanatory Exclusion and the Problem of Mental Causation'. *Information, Semantics and Epistemology*, ed. Enrique Villanueva. Oxford: Blackwell. pp.36-56.
- 1992: ' "Downward Causation" in Emergentism and Non-Reductive Physicalism', in Beckermann, A. et al. (1992).
- 1993a: *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press.
- 1993b: 'Causation, nomic subsumption, and the concept of event', in Kim, J. (1993a).
- 1993c: 'Events as property exemplifications', in Kim, J. (1993a).
- 1993d: 'Epiphenomenal and Supervenient Causation', in Kim, J. (1993a).
- 1993e: 'Mechanism, purpose, and explanatory exclusion', in Kim, J. (1993a).
- 1993f: 'The myth of nonreductive materialism', in Kim, J. (1993a).
- 1993g: 'The nonreductivist's troubles with mental causation', in Kim, J. (1993a).
- 1993h: 'Can Supervenience and "Non-Strict Laws" Save Anomalous Monism?', in Heil and Mele (1993).
- 1993i: 'Concepts of supervenience', in Kim, J. (1993a).
- 1993j: 'Supervenience as a philosophical concept', in Kim, J. (1993a).
- 1993k: '"Strong" and "global" supervenience revisited', in Kim, J. (1993a).
- 1993l: ' Postscripts on supervenience', in Kim, J. (1993a).
- 1993m: 'Postscripts on mental causation', in Kim, J. (1993a).

1997a: 'Does the Problem of Mental Causation Generalize?' *Proceedings of the Aristotelian Society*, 97, pp. 281-297.

1997b: 'Supervenience, Emergence, and Realization in the Philosophy of Mind'. *Mindscapes: Philosophy, Science and the Mind* (eds.) M. Carrier & P.K. Machamer, pp. 271-93.

1999a: *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. (2nd Printing) Cambridge, Massachusetts: The MIT Press.

1999b: 'Making Sense of Emergence'. *Philosophical Studies*, 95, pp. 3-36.

Klagge, J. C. 'Supervenience: Model Theory or Metaphysics?' in Savellos and Yalcin (1995).

Kripke, S. A. 1980: *Naming and Necessity*. Oxford: Basil Blackwell.

Leibniz, G. 1898 *The Monadology*. trs. Latta, R. Oxford: Oxford University Press, 1696.

LePore, E. and Loewer, B. 1987: 'Mind Matters'. *Journal of Philosophy*, 84, pp.630-42.

LePore, E and McLaughlin, B. (eds.) 1985: *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell.

Lewis, D. 1973: 'Causation'. *Journal of Philosophy*, 70, pp. 556-67.

Lockwood, M. 1989: *Mind, Brain and the Quantum*. Oxford: Basil Blackwell.

Lombard, L. 1986: *Events: A Metaphysical Study*. London: Routledge and Kegan Paul.

London, F. and Bauer, E. 1983: 'The Theory of Observation in Quantum Mechanics', trans., in Wheeler and Zurek (1983).

Loux, M. J. 1978: *Substance and Attribute*. London: D Reidel.

Loux, M. J. (ed) 1976: *Universals and Particulars: Readings in Ontology*.

Notre Dame: University of Notre Dame.

Lowe, E. J. 1980: 'For Want of a Nail', *Analysis*, 40, pp. 50-2.

1989: *Kinds of Being: A Study of Individuation, Identity and the Logic of Sortal Terms*. Oxford: Basil Blackwell.

1993: 'The Causal Autonomy of the Mental'. *Mind*, 102, pp. 629-44.

1995: 'The Metaphysics of Abstract Objects'. *Journal of Philosophy*, 92, pp.509-24.

1996: *Subjects of Experience*. Cambridge: Cambridge University Press.

1998: *The Possibility of Metaphysics: Substance, Identity and Time*. Oxford: Clarendon Press.

1999a: 'Self, Agency and Mental Causation. *Journal of Consciousness Studies*, 6, pp. 225-239.

- 1999b: 'Abstraction, Properties, and Immanent Realism', in Rockmore, T. (ed) *The Proceedings of the Twentieth World Congress of Philosophy, Vol. 2, Metaphysics*, pp. 195- 205.
- 2000: 'Causal Closure Principles and Emergentism'. *Philosophy*, 75, pp. 571-585.
- 2000b: *An Introduction to the Philosophy of Mind*. Cambridge: Cambridge University Press.
- 2001a: 'The disappearance of mind—and body?' *American Journal of Psychology*, 114, p. 303-8.
- 2001b: 'Recent Advances in Metaphysics' (abstract). *Formal Ontology in Information Systems: Collected Papers from the Second International Conference*. Welty, C. & Smith, B. (eds.) New York: ACM Press.
- Forthcoming a: 'A defence of the four-category ontology'. *Argument und Analyse*. Moulines, C. U. & Niebergall, K. G. (eds.). Paderborn: Mentis-Verlag.
- Forthcoming b: 'Recent Advances in Metaphysics'.
- Macdonald, C. 1989: *Mind-Body Identity Theories*. London: Routledge & Kegan Paul.
- Macdonald, C. & G. 1986: 'Mental Causes and Explanation of Action'. *Philosophical Quarterly*, 36, pp. 145-58.
- 1991: 'Mental Causation and Non-Reductive Monism'. *Analysis*, 51, pp. 23- 31.
- McGinn, C. 1991: *The Problem of Consciousness*. Oxford: Basil Blackwell.
- 1997: 'Consciousness and Space' in Chalmers (1997).
- Mackie, J. L. 1980: *The Cement of the Universe*. Oxford: Oxford University Press.
- McLaughlin, B. 1989: 'Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical.' *Philosophical Perspectives*, 3, pp. 109-35.
- 1992: 'The Rise and Fall of British Emergentism' in Beckermann, A. et al. (1992).
- 1993: 'On Davidson's Response to the Charge of Epiphenomenalism', in Heil and Mele (1993).
- 1995: 'Varieties of Supervenience', in Savellos (1995).
- Malcolm, N. 1968: 'The Conceivability of Mechanism'. *Philosophical Review*, 77, pp. 45-72.
- Marras, A. 1994: 'Nonreductive Materialism and Mental Causation'. *Canadian Journal of Philosophy*, 24, pp. 465-493.
- Martin, C. B. 1980: 'Substance Substantiated'. *Australian Journal of Philosophy*, 58, pp. 3-10.
- 1993: 'The Need for Ontology: Some Choices'. *Philosophy*, 68, pp. 505-22.
- Mellor, D. H. 1991: *Matters of Metaphysics*. Cambridge: Cambridge University Press.

- Melnyk, A. 1997: 'How to Keep the "Physical" in Physicalism'. *The Journal of Philosophy*, 94, p. 622- 637.
- Miller, R. 1990: 'Supervenience Is a Two-Way Street', *Journal of Philosophy*, 87, pp. 695-701.
- Mills, E. 1996: 'Interactionism and Overdetermination'. *American Philosophical Quarterly*, 33, pp. 105-17.
- Moore, G. E. 1953: *Some Main Problems of Philosophy*. London: Allen & Unwin.
- Morgan, C. L. 1923: *Emergent Evolution*. London: Williams & Norgate.
- Neale, S. 1995: 'The Philosophical Significance of Gödel's Slingshot'. *Mind*, 104, pp. 761-825.
- Noordhof, P. 1998: 'Do Tropes Resolve The Problem of Mental Causation?' *Philosophical Quarterly*, 48, pp. 221-6.
- Oliver, A. 1996: 'The Metaphysics of Properties'. *Mind*, 105, pp. 1- 80.
- Olson, K. R. 1987: *An Essay on Facts*. Stanford: CSLI.
- Papineau, D. 1990: 'Why Supervenience'. *Analysis*, 50, pp. 66-71.
- 1991: 'The Reason Why'. *Analysis*, 51, pp. 37-40.
- 1993: *Philosophical Naturalism*. Oxford: Basil Blackwell.
- 1995: 'Arguments for Supervenience and Physical Realization', in Savellos and Yalcin (1995).
- 1998: 'Mind the Gap', in Tomberlin, J. (ed.), *Philosophical Perspective*, 12, pp. 373-88.
- 2000: 'The Rise of Physicalism' forthcoming in Stone, M. W. F. and Wolff, J. (2000)
- Papineau, D. and Spurrett, D. 1999: 'A Note on the Completeness of Physics'. *Analysis*, 59, pp. 25- 29.
- Penrose, R. 1989: *The Emperor's New Mind*. Oxford: Oxford University Press.
- Poland, J. 1994: *Physicalism: The Philosophical Foundationalism*. Oxford: Clarendon Press.
- Post, J. F. 1995: ' "Global" Supervenient Determination: Too Permissive?' in Savellos (1995)
- Putnam, H. 1975: 'The Meaning of "Meaning"'. *Mind, Language and Reality: Philosophical Papers Volume 2*. Cambridge: Cambridge University Press.
- 1980: 'The nature of mental states', in Block (1980). Vol. 1.
- 1982: 'Why there isn't a ready-made world'. *Synthese*, 51, pp. 141-68.

- Quine, W. 1960: *Word and Object*. Cambridge: MA: MIT Press.
- Quinn, P. 1997: 'Tiny Selves: Chisholm on the Simplicity of the Soul', in Hahn (1997).
- Robb, D. 1997: 'The Properties of Mental Causation'. *The Philosophical Quarterly*, 47, pp. 178-94.
- 2001: 'Reply to Noordhof on Mental Causation'. *The Philosophical Quarterly*, pp. 92-93.
- Russell, B. 1918: *The Philosophy of Logical Atomism*. Reprinted in *Russell's Logical Atomism*, ed. D. Pears. London: Fontana/ Collins, 1972.
- Salmon, W. 1984: *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- 1994: 'Causality without Counterfactuals'. *Philosophy of Science*, 61, pp. 297-312.
- Savellos, E. & Yalcin, U. (eds.) 1995: *Supervenience: New Essays*. Cambridge: Cambridge University Press.
- Seager, W. 1991: *Metaphysics of Consciousness*. London: Routledge.
- Searle, J. R. 1984: *Minds, Brains and Science*. London: BBC.
- 1992: *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
- Segal, G. & Sober, E. 1991: 'The Causal Efficacy of Content'. *Philosophical Studies*, 63, pp. 1-30.
- Shear, J. 1995: *Explaining Consciousness- The 'Hard Problem'*. Cambridge, MA: MIT Press.
- Shimony, A. 1993: *Search for a Naturalistic World View*. Vol. 2. Cambridge: Cambridge University Press.
- Shoemaker, S. 1984a: *Identity, Cause and Mind: Philosophical Essays*. Cambridge: Cambridge University Press.
- 1984b: 'Causality and Properties' in Shoemaker, S. (1984a).
- 1984c: 'Identity, Properties and Causality' in Shoemaker, S. (1984a).
- Simons, P. 1987: *Parts: A Study in Ontology*. Oxford: Clarendon Press.
- Sklar, L. 1992: *Philosophy of Physics*. Oxford: Oxford University Press.
- Smart, J. J. C. 1978: 'The Content of Physicalism'. *Philosophical Quarterly*, 28, pp. 339-341.
- Smith, P. 1984a: 'Bad News for Anomalous Monism'. *Analysis*, 42, pp. 220-225.
- 1984b: 'Anomalous Monism and Epiphenomenalism: A Reply to Honderich'. *Analysis*, 44, pp. 83-6.
- Smith, P. & Jones, O. 1986: *The Philosophy of Mind: An Introduction*. Cambridge:

Cambridge University Press.

- Snowdon, P. 1989: 'On Formulating Materialism and Dualism', in Heil, J. (ed) *Cause, Mind and Reality: Essays in Honour of C. B. Martin*. Dordrecht: Kluwer Press.
- Sosa, E. 1993: 'Davidson's Thinking Causes' Heil and Mele (1993).
- Sosa, E. & Tooley, M. (eds.) 1993: *Causation*. Oxford: Oxford University Press.
- Sperry, R. W. 1952: 'Neurology and the Mind-Brain Problem'. *American Scientist*, 40, pp. 291-312.
- Stapp, H. P. 1997: 'The Hard Problem: A Quantum Approach', in Chalmers (1997).
- Stone, M. W. F. & Wolff, J. (ed.) 2000: *The Proper Ambition of Science*. New York: Routledge.
- Sturgeon, S. 1998: 'Physicalism and Overdetermination'. *Mind*, 110, pp. 411- 432.
- Taylor, B. 1985: *Modes of Occurrence*. Oxford: Basil Blackwell.
- Teller, P. 1992: 'A Contemporary Look at Emergence', in Beckermann, A. et al. (eds.)
- Tooley, M. 1987: *Causation: A Realist Approach*. Oxford: Clarendon Press.
- Tye, M. 1982: 'A Note on the Synonymy Principle of Property Identity'. *Analysis*, 42, pp. 52-55.
- Warner, R. and Szubka, T. (eds.). 1994: *The Mind-Body Problem* (Oxford: Blackwell).
- Weinberg, S. 1999: 'A Unified Physics by 2050'. *Scientific American*.
- Wheeler, J. A. and Zurek, W. H. (eds.) 1983: *Quantum Theory and Measurement*. Princeton: Princeton University Press.
- Wigner, E. P. 1961: 'Remarks on the Mind-Body Problem', in Wheeler and Zurek (1983)
- Williams, D. C. 1966: *Principles of Empirical Realism*. Springfield: Charles C. Thomas.
- Wittgenstein, L. 1922: *Tractatus Logico-Philosophicus*. Trans. Ck K. Ogden. London: Routledge & Kegan Paul.
- Woolhouse, R. 1985: 'Leibniz's Reaction to Cartesian Interaction'. *Proceedings of the Aristotelian Society*, 86, pp. 69-82.
- Yablo, S. 1992: 'Mental Causation'. *Philosophical Review*, 101, pp. 245-280.

